ISTANBUL TECHNICAL UNIVERSITY ★ GRADUATE SCHOOL

DESIGNING AN EXPERT SYSTEM FOR NON-EXPERT USERS IN ORAL HEALTH: STORY OF A HYBRID DESIGN RESEARCH

M.Sc. THESIS

Deniz GÖÇHAN

Department of Industrial Design

Industrial Design Program

JUNE 2023

ISTANBUL TECHNICAL UNIVERSITY ★ GRADUATE SCHOOL

DESIGNING AN EXPERT SYSTEM FOR NON-EXPERT USERS IN ORAL HEALTH: STORY OF A HYBRID DESIGN RESEARCH

M.Sc. THESIS

Deniz GÖÇHAN
(502171903)

Department of Industrial Design

Industrial Design Program

Thesis Advisor: Prof. Dr. H. Hümanur BAĞLI

JUNE 2023

İSTANBUL TEKNİK ÜNİVERSİTESİ ★ LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ

AĞIZ SAĞLIĞI ALANINDA UZMAN OLMAYAN KULLANICILAR İÇİN UZMAN SİSTEM TASARLAMAK: HİBRİT TASARIM ARAŞTIRMASI HİKAYESİ

YÜKSEK LİSANS TEZİ

Deniz GÖÇHAN
(502171903)

Endüstriyel Tasarım Anabilim Dalı

Endüstriyel Tasarım Programı

Tez Danışmanı: Prof. Dr. H. Hümanur BAĞLI

HAZİRAN 2023

Deniz GÖÇHAN, an M.Sc. student of İTU Graduate School student ID 502171903, successfully defended the thesis/dissertation entitled "DESIGNING AN EXPERT SYSTEM FOR NON-EXPERT USERS IN ORAL HEALTH: STORY OF A HYBRID DESIGN RESEARCH", which he prepared after fulfilling the requirements specified in the associated legislations, before the jury whose signatures are below.

Thesis Advisor: **Prof. Dr. H. Hümanur BAĞLI** ....
Istanbul Technical University

Jury Members: **Prof. Dr. Şebnem TİMUR** ....
Istanbul Technical University

Jury Members: **Assoc. Prof. Dr. A. Selami ÇİFTER** ....
Mimar Sinan Fine Arts University

Date of Submission : 24 May 2023
Date of Defense : 23 June 2023

To Alan Turing,
***the founding father of artificial intelligence who saved millions of lives*,**

FOREWORD

The master's education and this thesis represent much more than just a degree and research in the field for me; rather a turning point in my life to succeed in studying in this school, which was my childhood dream.

Design education has changed my personality and my perspective on life. Besides design, I learned a lot about myself and what I am capable of. I grew up. After years of training in the natural sciences as a dentist, I enjoyed the freedom that social studies give a researcher. I also gained a deeper understanding of the other values that make people human, besides flesh and bone. This education has a great role in the emergence of the project that I have been working on for the past three years, aiming to develop artificial intelligence in pediatric oral health.

I would like to thank my teacher, mentor, and dearest friend, Prof. Dr. H. Hümanur BAĞLI. I learned a lot from her about not only design but also friendship. She is the closest witness to the change in me and one of the main characters of this story. I am grateful for her sincerity, love, and company.

In 2015, when I knew nothing but wanted to study design, the first person I contacted in the department was Prof. Dr. Şebnem TİMUR. She welcomed me wholeheartedly, and thanks to her guidance, I was able to get accepted to the school. I am also grateful for everything I learned from her.

I would like to thank all the professors and friends who taught me something on this path and made me who I am today.

I would like to thank TÜBİTAK for the financial support. Thanks to the grant, this project has become my job and changed the course of my life.

I would like to express my gratitude towards Nuray Öner GÜCİN and Elçin KURBANOĞLU for their invaluable support. Their presence throughout the entire process was greatly appreciated and meant a lot.

I cannot thank my mother, Sevim GÖÇHAN, and my father, Fahri GÖÇHAN enough for their financial and moral support throughout the project. Without them, this project would not have been the subject of this thesis.

Sometimes all you need is people who see your potential and believe in you. I am very happy and feel lucky that life has matched me up with these people. So, thank you life…

June 2023 Dr. Deniz Göçhan

TABLE OF CONTENTS

Page

ABBREVIATIONS

AAPD	: American Academy of Pediatric Dentistry
AI	: Artificial Intelligence
AIC	: Acıbadem Incubation Center
AIMDD	: Active Implantable Medical Device Directive
CE	: Conformité Européenne
DRS	: Dental Recording System
EG	: Emergency Guide
EU	: European Union
FDA	: United States Food and Drug Administration
GDPR	: General Data Protection Regulation
HIPAA	: Health Insurance Portability and Accountability Act of 1996
HONcode	: Health on the Net Foundation Code of Conduct
IOS	: iPhone Operating System
MARS	: Mobile App Rating Scale
MDD	: Medical Device Directive
MDR	: Medical Device Regulation
ML	: Machine Learning
ND	: Nearby Dentists
NHS	: National Health Service of the United Kingdom
NLP	: Natural Language Processing
OEG	: Oral Examination Guide
OHPP	: Oral Hygiene Product Platform
OHTS	: Oral Hygiene Tracking System

RfD : Research for Design

RiD : Research into Design

RtD : Research through Design

SaMD : Software as a Medical Device

SAS : Symptom Assessment System

SBOSC : Scientific Basis of Oral Self-care

SSV1 : Symptom System Version 1

SSV2 : Symptom System Version 2

TMMDA : Turkish Medicines and Medical Devices Agency

TÜBİTAK : Scientific and Technological Research Council of Turkiye

UDI : Unique Device Identification

UI : User Interface

UX : User Experience

WHO : World Health Organization

LIST OF TABLES

Page

LIST OF FIGURES

DESIGNING AN EXPERT SYSTEM FOR NON-EXPERT USERS IN ORAL HEALTH: STORY OF A HYBRID DESIGN RESEARCH

SUMMARY

The subject of this thesis is the research, development, and design process of an expert system that I designed for non-experts in the field of pediatric oral health, which started as an enterprise project and later received the TÜBİTAK 1512 Techno-initiative Capital Support Program grant. In this thesis, you will read about the maturation of my idea that I created as a dentist, its transformation into a mobile application prototype within the scope of the TÜBİTAK project, the testing processes, the difficulties encountered, and how I tried to overcome them.

In empirical research, the independence of the researcher and the notion that examines the subject from an outsider's perspective are very important. I, on the other hand, was the leading actor in the three-year enterprise story that I addressed as a researcher in this thesis. At every stage of the product from an idea to a prototype; I spent a lot of time applying to institutions, developing the algorithm, and designing the interfaces and the system. For these reasons, I did not want to write down the research I conducted in the process as if I were an observer. Because that would be quite wrong. Thus, the complexity of the process and my inseparable relationship with the product led to the need to adopt a hybrid approach to the subject.

In my approach, I combined the three forms of design research that Frayling described. The "Research through Design" approach allowed me the freedom to convey what I experienced in the process from my perspective, without skipping any steps. Since it is not possible to generalize about entrepreneurship based on my own story, there could not be a more appropriate method in which I would be involved as a subject. The research I conducted to enable non-expert users to use an expert system was exactly in line with the "Research for Design" approach. Because at the end of the day, my goal was to improve the product. The "Research into design" approach enabled me to document data that emerged independently of the product I designed and could contribute to the literature. I am aware that in this approach I have adopted, a portrayal of dissociative identity emerges. I became a "Research through Design" researcher while describing my process, a designer when researching to improve the product, and sometimes a design researcher who intends to contribute to the literature. A researcher can wear these hats in different periods. I wore all these hats during one study. Therefore, in this thesis, you will witness hybridity just like my process.

This thesis touches on more than one area. Therefore, what I have told can be evaluated from many different angles. It can be considered as the process of developing an entrepreneur's product, the place of digital products in the field of oral health, a critique of the way dentists handle these products in the literature, the difficulties in developing a medical device, designing of artificial intelligence and

human interaction, the effort to bring an expert system closer to the non-expert user, the design’s role in all these aspects or as a different method attempt at design.

AĞIZ SAĞLIĞI ALANINDA UZMAN OLMAYAN KULLANICILAR İÇİN UZMAN SİSTEM TASARLAMAK: HİBRİT TASARIM ARAŞTIRMASI HİKAYESİ

ÖZET

Bu tez, bir girişimcilik fikri olarak başlattığım ve sonrasında "Çocuk Ağız ve Diş Sağlığı Alanında Bulanık Mantık Yönteminin Kullanılmasıyla Uzman Sistemin Ön Prototipinin Geliştirilmesi ve Ticarileşme Amacıyla Mobil Uygulamaya Uyarlanması" isimiyle TÜBİTAK 1512 Tekno-girişim Sermaye Destek Programı hibesini almaya hak kazanan bir inovasyon projesini konu edinmektedir. Bu projede, uzman olmayan kullanıcılara yani ebeveynlere çocuk ağız sağlığı alanında olası tanıyı sunan bir uzman sistem prototipini geliştirmeyi amaçladım. Dolayısıyla tezde, bir diş hekimi olarak ortaya koyduğum fikrin olgunlaşmasını, TÜBİTAK projesi kapsamında bir mobil uygulama prototipine dönüşmesini, Ar-Ge sürecinde yapılan tasarım araştırmalarını, test edilme süreçlerini, tüm bu süreçlerde karşılaşılan zorlukları ve bu zorlukların nasıl üstesinden gelmeye çalıştığımı okuyacaksınız.

Girişimcilik/inovasyon sürecimde yaptığım tasarım araştırmalarını tez kapsamında ele almaya karar verdiğimde çeşitli sorunlarla karşılaştım. Bu sorunlardan birincisi bir araştırmacı olarak süreçten ve üründen bağımsız olmamamdı. Ürünü geliştiren ve tasarlayan kişiydim. Araştırmaları geliştirdiğim ürünü iyileştirmek amacıyla yapmıştım. Bu araştırmaları girişimcilik sürecinden bağımsız incelemek de anlamsız olacaktı çünkü içeriği özgün olsa bile yöntemsel açıdan tasarım süreçlerinde sıklıkla yapılan araştırmalardı. Ayrıca bu sürecin tez konum olmasına, sürecin ortasında karar vermiştim. Ampirik araştırmalarda, araştırmacının bağımsızlığı, konuyu dışarıdan bir gözle irdeleyen bakış açısı oldukça önemlidir. Araştırma önden planlanır; bir hipotez ortaya atılır ya da bir soru sorulur ve hipotez test edilir ya da bir cevap aranır. İçinde bulunduğum durumun karmaşıklığı ve bir araştırmacı olarak çoklu pozisyonum bu kuralların hiçbirine uymuyordu. Ben zaten genellenebilir, yeniden üretilebilir bir bilgi üretmiyordum. Sadece girişimcilik hikayemi ve Ar-Ge sürecimi, bir araştırmacı olarak pozisyonumu ve amaçlarımı açıkça belirterek tasarım çerçevesinden anlatmak istiyordum. Dolayısıyla, sürecin karmaşıklığı ve ürün ile ayrılamaz ilişkim, tez için hibrit bir yaklaşımı benimseme ihtiyacını doğurdu.

Yaklaşımımda Frayling'in tanımladığı üç tasarım araştırması biçimini birleştirdim. "Tasarım yoluyla araştırma" yaklaşımı bana herhangi bir aşamayı atlamadan kendi gözümden süreci aktarma özgürlüğünü tanıdı. Kendi hikayemden yola çıkarak girişimciliğe dair bir genelleme yapılamayacağı için özne olarak benim de yer alacağım daha uygun bir yöntem olamazdı. Bu şekilde fikrin ortaya çıkmasından, girişimin geldiği son noktaya kadar yapılan tüm işleri ve aldığım kararları ben diliyle yazabildim.

Uzman bir sistemi uzman olmayan kullanıcının kullanabilmesi için yaptığım araştırmalar tam olarak "Tasarım için araştırma" yaklaşımına uyuyordu. Çünkü

günün sonunda süreç içindeki amacım tasarladığım ürünü iyileştirmekti. Bu amaç doğrultusunda toplam üç tasarım araştırması yaptım. Çocuklarda ağız hastalıklarının olası tanısını koyacak uzman sistemin ilk prototipini tamamladığımda önemli bir engelle karşılaştım. Uzman terimleri uzman olmayan kullanıcılar nasıl anlayacaklardı? Çünkü kullanıcının semptom arama kutusuna ne yazdığını anlayacak "Doğal Dil İşleme" kabiliyeti algoritmada mevcut değildi. Bu nedenle, onların ağız hastalıklarını ve bu hastalıklara ait belirtileri nasıl tanımladığını öğrenmeye karar verdim. Birinci araştırmada, beş uzman olmayan katılımcıya internetten bulduğum kırk ağız hastalığı fotoğrafı gösterdim ve burada gördükleri hastalıkları sanki telefonda bir diş hekimine anlatır gibi tarif etmelerini istedim. Elde edilen veriler ışığında semptom tanımlarını ve bağlantılı olarak sınıflandırmaları ve algoritmayı değiştirdim. Bütün semptomları dokuz alt başlık altında topladım. Yeni semptom sistemine uygun olarak kullanıcıların ilgili semptomları seçtikleri arayüzleri Figma'da tasarladım ve ikinci tasarım araştırması için iki farklı kullanıcı akışına göre iki prototip oluşturdum. İkinci araştırmada, birinci araştırmadaki katılımcıların bu arayüzleri kullanıp kullanamayacaklarını test ettim. Onlara çocuk ağız hastalıklarıyla ilgili toplam sekiz vaka hikayesi okudum ve ağız içi fotoğraflar gösterdim. Katılımcılardan bu fotoğraflarda gördükleri ve hikayede tespit ettikleri belirtileri göz önüne alarak iki farklı prototipi kullanmalarını istedim ve görüşmeleri kaydettim. Kullanıcıların büyük bir kısmı yeni oluşturulan semptom sistemini arayüzler üzerinde kullanabildiler. Sonuçların tatmin edici çıkması ile bir sonraki aşama olan kodlama sürecine geçebildim. Algoritma ve arayüzler, Android Studio kullanılarak yazılımcı Mehmet Demir tarafından mobil uygulama olarak kodlandı. Üçüncü araştırmada mobil uygulama formunda geliştirdiğimiz uzman sistemi üç farklı kullanıcı grubuna, ebeveynler (N:3), tasarımcılar (N:3) ve diş hekimlerine (N:3) test ettirdim. Bu araştırmada katılımcılardan, oluşturduğum bir kurgu vaka hikayesindeki verileri sanki kendi çocuklarının verileriymiş gibi kabul etmelerini, çocuk için bir profil oluşturmalarını, hikayedeki belirtileri sisteme girerek bir değerlendirme yapmalarını ve bu faaliyetler sırasında sesli düşünmelerini istedim. Araştırma esnasında hem ekran görüntüsü hem de ses kaydı aldım ve sonuçları içerik analizi yaparak değerlendirdim.

Üçüncü araştırmada ürünün kendisinden bağımsız olarak başka sonuçlar da elde ettim. Genelde kullanılabilirlik araştırmaları son kullanıcının dahil edilmesiyle gerçekleştirilir. Ben ise bu araştırmada, ürünün temel aldığı iki alandaki uzmanları; tasarımcıları ve diş hekimlerini de dahil ettim. Ancak bu farklı katılımcıların değerlendirme esnasında hangi noktaların üzerinde daha çok durduklarını ve bu yeni yöntemin verileri hangi açılardan zenginleştirdiklerini sunmak "Tasarım için araştırma" yaklaşımına uygun değildi. Araştırmada elde edilen aynı sonuçları farklı bir bağlamda değerlendirmek, farklı bir yaklaşım gerektiriyordu. Bu nedenle, bu değerlendirmeyi Frayling'in son sınıfı olan "Tasarım adına araştırma" yaklaşımını benimseyerek yaptım.

Benimsediğim bu yaklaşımda ortaya bir kişilik bölünmesi tablosunun çıktığının farkındayım. Sürecimi anlatırken bir eylem araştırmacısı, ürünü iyileştirmek için araştırırken bir tasarımcı ve bazen de literatüre katkı yapmak isteyen bir tasarım araştırmacısı oldum. Bir araştırmacı farklı zaman dilimlerinde bu şapkaları takabilir. Ben ise bu şapkaların hepsini bir çalışma başlığı altında taktım. Çünkü mevcut yöntemlerin hiçbiri bu kompleks süreci anlatmak için yeterli değildi. Sürecin ve bir araştırmacı olarak konumumun melezliği benimsediğim yaklaşıma sirayet etti. Bu nedenle Frayling'in sınıflandırmasındaki bütün tasarım araştırması türlerini bir araya getirerek sürecimi anlattım.

İnovasyon projemdeki tasarım sürecini merkeze koyduğum bu tezde, ürünün temas ettiği dört alanda (İnsan- Bilgisayar Etkileşimi, Uzman Sistemler, Ağız Sağlığı/Mobil Uygulamalar ve Tıbbi Cihazlar) kapsamlı bir literatür incelemesi de gerçekleştirdim. Literatür incelemesini iki çatı altında topladım: İnsan-bilgisayar etkileşimi ve ağız sağlığı mobil uygulamaları/medikal cihazlar.

Birinci kısımın ilk bölümünde insanı yani kullanıcıyı mevcut literatürün nasıl sınıflandırdığını anlattım ve bu sınıflandırmaların eksikliklerini ortaya koyarak uzman sistemler ya da bir uzmanlık içeren sistemler bağlamında yeni bir kullanıcı sınıflandırması önerisinde bulundum. Bu sınıflandırmada uzman sistemin içerdiği uzmanlığı referans noktası alarak kullanıcıları uzman ve uzman olmayan kullanıcılar olarak ikiye ayırdım. Bu iki grubu bilgi, deneyim ve kabiliyet olmak üzere üç düzlemde alt gruplara ayırdım İkinci bölümde ise uzman sistemlerin ne olduğunu, hangi komponentlerden oluştuğunu örnekler vererek açıkladım. Üçüncü bölümde etkileşimi etkileyen parametreleri, insan-bilgisayar etkileşimi ve tasarım literatüründe hangi açılardan incelediklerini değerlendirdim. Bu değerlendirme sonucunda, kullanıcı profillemesinde yardımcı olabilecek üç boyutlu bir model önerisinde bulundum. Bu modeli kullanıcı sınıflandırmasıyla ilişkilendirerek açıkladım.

İkinci kısımda, ağız sağlığı uygulamalarını klinik ve tasarım açılarından değerlendiren çalışmaları ve geliştirilme aşamalarını anlatan çalışmalarda kullanılan tasarım yöntemlerini ele aldım. Literatürdeki çalışmalarda var olan yöntemsel eksiklere değinerek eleştiri ve önerilerde bulundum. Ağız sağlığı uygulamalarını tıbbi cihaz kapsamında değerlendiren yönetmelikleri, uyum ve sertifika süreçlerini inceledim. Bir tıbbi cihaz üreticisi olarak yönetmeliklerin sektöre getireceği fayda ve riskler üzerine yorumlarda bulundum.

Bu tezde, ağız sağlığı alanında uzman olmayan kullanıcılar için bir uzman sistem geliştirdiğim proje sürecini hibrit bir araştırma yöntemi benimseyerek anlattım. Buna ek olarak, ürünün temas ettiği alanlarda kapsamlı bir literatür incelemesi yaptım ve bu alanlara katkı sağlamak amacıyla eleştirilerde ve önerilerde bulundum. Öneri, eleştiri ve yorumlarımın yer aldığı bölümleri, literatür incelemesi ve süreçte elde ettiğim deneyimlerin bir çıktısı ya da bir girişimci, diş hekimi, tıbbi mühendis, tasarımcı ve araştırmacının öz-yansıtması olarak kabul edebilirsiniz. Bu çıktılar çok farklı açılardan, sağlık alanında bir girişimcilik hikayesi, yapay zekâ ve insan etkileşiminin tasarlanması, uzman bir sistemi uzman olmayan kullanıcıya yaklaştırma çabası, tasarımın tüm bu süreçlerdeki rolü ya da tasarımda farklı bir yöntem denemesi olarak ele alınabilir. Bu nedenle bu tezin, mobil uygulamalar, uzman sistemler ya da bir uzmanlık içeren sistemler üzerine çalışan girişimciler, sağlık uzmanları, kural koyucular, tıbbi cihaz üreticileri, mühendisler, tasarımcılar, akademisyenler ve araştırmacılara faydalı olabileceğine inanıyorum.

1. INTRODUCTION

1.1 Reflexive Prelude

Since my approach in this thesis is strongly influenced by the principles of action research and “Research through Design” is a part of my hybrid method, I need to briefly introduce myself, my position, and my perspective on life. Thus, you can more easily understand my place and role in this study, the aspects of my criticisms, and comments, and especially how I handle the design process part in which I speak in the first person.

As a dentist, it would be right to start by answering the question of how my path crossed with industrial design. A very simple answer that you would often encounter in Turkey: The major I always wanted to study as a kid was industrial design. Since my score in those years was not enough for the industrial design major at Istanbul Technical University (ITU) and Middle East Technical University, and since I was raised in a middle-class family by teacher-parents, I chose dentistry as a profession where I would not have "future anxiety".

I graduated from Ege University Faculty of Dentistry in 2008. Although my bond with the profession has been shaken at times, I can say that I mostly enjoyed being a dentist or I never hated it. In 2009, I started my doctorate in the Department of Periodontology at Başkent University and studied it with pleasure as well. My area of interest during my Ph.D. was oral medicine and I made publications in which I generally reported the findings of systemic or genetic diseases in the mouth. Diagnosing oral pathologies was my favorite thing to do because it always felt like solving a puzzle. After completing my doctorate in 2013, I moved to Istanbul and started to work as an oral surgeon at Istanbul Aydın University Dental Hospital. I worked there full-time for about three years. Although my income was good, my job started to not satisfy me after a while. Something was missing in what I was doing for a living and I needed to move to a field where I could use my creativity.

While trying to figure out what to do in this quest, I started taking drawing lessons in 2015. In the first lesson, my art teacher, Fatma Demirtaş, asked me if I ever wanted to be an industrial designer because she thought my drawings were very neat. This was quite shocking and hit the bull’s eyes. It reminded me of the fact that I had a childhood dream. During our sessions, she motivated me so hard to take the aptitude test, so I suddenly found myself preparing for the university exam again. However, I knew in my heart that I was not ready to study undergraduate again, and I was also aware that it was not enough to take a drawing course for 6 months to get accepted. However, I took the aptitude test anyway and of course, I did not succeed. Then I decided that doing a master's in design was a more realistic plan. I sent an e-mail to Prof. Şebnem Timur at ITU and asked for an appointment. In our interview, she told me that dentistry should have been on the list of eligible degrees so that I could apply for the master’s. Thanks to her, I wrote a petition to the board of directors for the addition of dentistry to the list and my request was accepted.

In the fall of 2016, I took the Design Thinking course given by Prof. Hümanur Bağlı as a guest student to get to know the field and see what I could do. Students who took this master's course were expected to improve the projects of the startups they were matched with by using design methods and to bring a designer perspective to the entrepreneurs. Luckily, I was matched with Erhan Ermek. Erhan was developing a POCT device that measures clotting time. For his project, I did interviews with different healthcare professionals and learned what they cared about in a POCT device. In addition, I evaluated the perceptions of designers and 60 diabetic patients by showing them ten different blood glucose meter POCT device designs. I used several semantic differential scales such as trust, seriousness, professionalism, quality, etc. to understand which design elements came into effect in such parameters. I likened Design Thinking to medicine in terms of diagnosing problems in projects using design methods and making an intervention based on the results obtained. Thanks to this course, it was easy for me to enter the world of design.

In the spring term of 2017, I applied for the master's program at ITU but was not accepted. Encouraged by my success in the previous course, I took three more courses as a guest student: Design Semantics, Paradigms of Design, and Theoretical Studies in Design. In the Design Semantics course, I conducted user research that evaluated toothbrush designs semantically. In May, I presented my final paper for the

Design Thinking course at the 4T Design and Design History Society's Cross-Abilities Symposium. Meanwhile, Erhan offered to work together because he liked what I had done for his project. When his venture was accepted by Bayer's Grants4Apps program, we went to Berlin in June for investor meetings. It was an incredible experience for me. Over the course of a two-year collaboration with Erhan, I had the opportunity to take responsibility for the application process for numerous incubation centers and pitch events. Within this role, I was tasked with managing both the medical background and design aspects of the project. Through this experience, I gained a wealth of knowledge surrounding the intricacies of entrepreneurship and the many challenges that come along with it.

In the following year, I applied again for the master's program and became the first dentist admitted to the department. After a one-term scientific preparation process, I took other post-graduate courses. Particularly, the Ethnographic Methods in User-Centered Design and Cultural Approaches to Design courses made me fall in love with social and cultural studies. For the Cultural Approaches to Design course, I wrote a paper exploring the experience of users from different cultures with different types of bidets. After years of positivist education, design education taught me to look at research forms, methods of acquiring knowledge, and even medicine from a different perspective. For this reason, besides the happiness of fulfilling my childhood dream, I see this process as one of the most significant turning points of my life.

My ability to develop artificial intelligence for parents in the field of pediatric oral health has been realized thanks to the comprehensive knowledge about the field and most importantly the scientific notion I gained through dentistry education, as well as the designer perspective, qualitative research methods, production processes, user research I learned through design education. This innovation is the fruit of my 19 years of university education, and I hope it will be of some use to humanity.

You can find some of my thoughts related to the entrepreneurial parts of my story in Chapter 3, where I describe my design process. In this thesis, you will encounter many of my identities: a dentist, an inventor, a medical device manufacturer, a medical engineer, a design student, an entrepreneur, and above all, a researcher.

1.2 Structure of the Thesis

This thesis is the outcome of my entrepreneurship project where I designed an expert system in the field of pediatric dentistry for non-expert users. It was quite challenging to deal with such a multi-layered and multi-disciplinary project within the scope of a master's thesis in design. Because this three-year project is located right at the intersection of four different areas: Oral health & mobile apps -if considered one-, medical devices, expert systems, and human-computer interaction (HCI). Although I tried to review relevant literature on these four fields from a design perspective as much as possible, some parts that were out of the scope of the design remained in the thesis for a better understanding of the technology I have developed. For this reason, sections such as the clinical approach to oral health mobile applications, expert systems, or regulations governing medical devices can be considered subsections that support the main framework.

The thesis is divided into 5 chapters:

Chapter 1 is an introductory part where I review the literature on four fields in which the product intersects. It consists of two main sub-sections: (1) HCI in expert systems and (2) oral health apps.

Under HCI, there are three headings: “Defining users”, “Introduction to AI, and expert systems” and “Designing interaction”. In the “Defining users” section, after reviewing the current definitions in the literature, I define and categorize possible end users of an expert system and their characteristics. The “Introduction to AI and expert systems” briefly introduces the technical aspects of expert systems and their components. “Designing interaction” focuses on the factors that affect the interaction between expert systems and end users. It is a section where I mention semiotics, affordances, user models, and profiling, and I propose a new model for user profiling.

The literature in the oral health apps section is divided into two. In the first part, how oral health mobile apps are evaluated in the mDentistry literature is discussed. In the “Clinical approach in mDentistry” section, after a brief introduction to both the health and socio-economic consequences of oral diseases, it is focused on what intervention purposes oral health apps are used for and how their performance is evaluated in the mDentistry literature. “The categorizations of oral health apps in the literature” is a section that focuses on how existing oral health apps on the market are classified in the literature. “The design approach in the evaluations by health experts” reviews how the design and usability of oral health apps are evaluated by health experts. “The design methods adopted in the development processes of oral health apps” emphasizes the design methods adopted by the developers in the development processes of oral health applications. “The issues of approaches in the mDentistry literature” discusses the methodological problems and drawbacks of the studies in the mDentistry literature. In the second part, the unique dynamics of medical devices are addressed such as governing regulations for medical apps, the role of designers in the field, and the prospects and challenges. The “Governing regulations for medical apps” section focuses on the governing rules in FDA and MDR regulations and compliance processes regarding health apps. “The role of designers in the medical device development processes” emphasizes the challenges of medical device development processes, the role of designers, and how the design service is perceived by the manufacturers. The “Prospects and challenges” section discusses the potential benefits and challenges that the newly enacted regulations to which health apps are subject will bring to the field.

Chapter 2 is the part where I explain how I approach my design process within the research context. In the beginning, I introduce action research and its principles, then Frayling’s (1994) classification of design research, and finally I elucidate the hybrid research approach I created by combining Frayling's classification.

Chapter 3 is devoted to a meta-reflection of the design process of the expert system I developed for pediatric oral health. In this part, the development process is elaborated considering its stages. Research and development activities, especially design research carried out from the idea to the test stage, are explained with an active voice in this section.

Chapter 4 is the part where I discuss the research approach that I have adopted in this thesis and my design process.

Chapter 5 is the section where I summarize my findings, critiques, proposals, and comments as conclusive remarks.

This thesis has some structural differences from the theses written at Istanbul Technical University and even those we encounter frequently in academia. In the four subsections of the literature review ("Redefining the user in the context of expert systems", "A model proposition for profiling the users of expert systems", "The issues of approaches in the mDentistry literature" and "Prospects and challenges of regulations"), you will read my suggestions, criticisms, and comments on the relevant subjects. According to the school's thesis writing style, these should have been included in the discussion. However, it would not be right to place these subsections there where my design process was mostly discussed. Because they were closely related to the sections in the literature review. Placing them in the discussion would cause a break in the flow.

This unusual structure carries some risks of misunderstanding. Since the reader will read these sections first, they can naturally expect to see a reflection of my suggestions and criticisms in my design process whereas they have emerged as a result of my experiences in the process and extensive literature review. If I had been able to write the thesis in chronological order, you would have read my design process first and then the literature review and hence my suggestions and criticisms. Unfortunately, the current thesis template of the school did not allow for the eclectic nature of the subject that this thesis was dealing with. Therefore, I would like you to accept the first part, the literature review, both as an introduction to the areas in which the project intersects and as an output that has been put forward based on the experience and data I have obtained at the end of the process.

While active voice is used in the sections where the product's design process is explained, passive voice is preferred in the sections where approaches or relevant studies in the literature are discussed. Both writing styles are used in the sections such as the discussion or research approach where the process and the literature are discussed together. Figure 1.1 shows the writing styles generally adopted in the chapters.

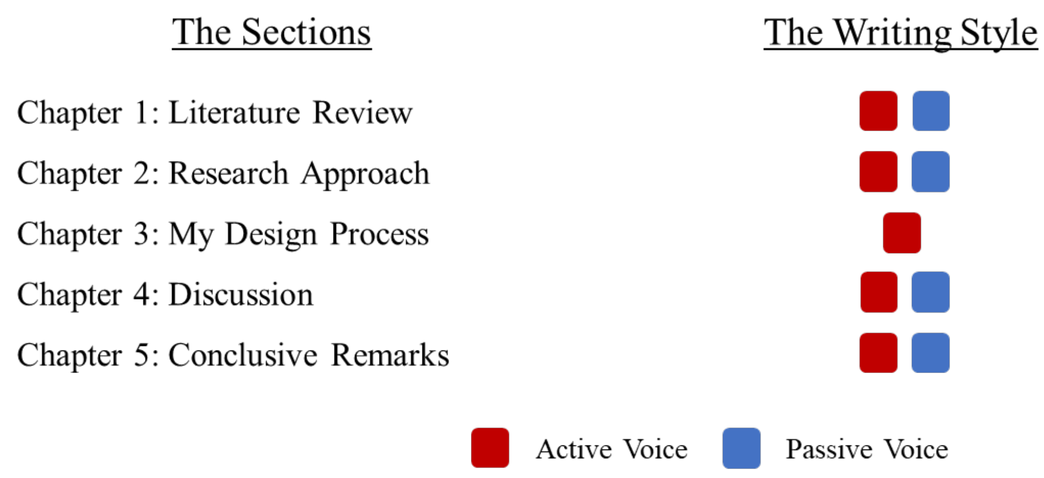


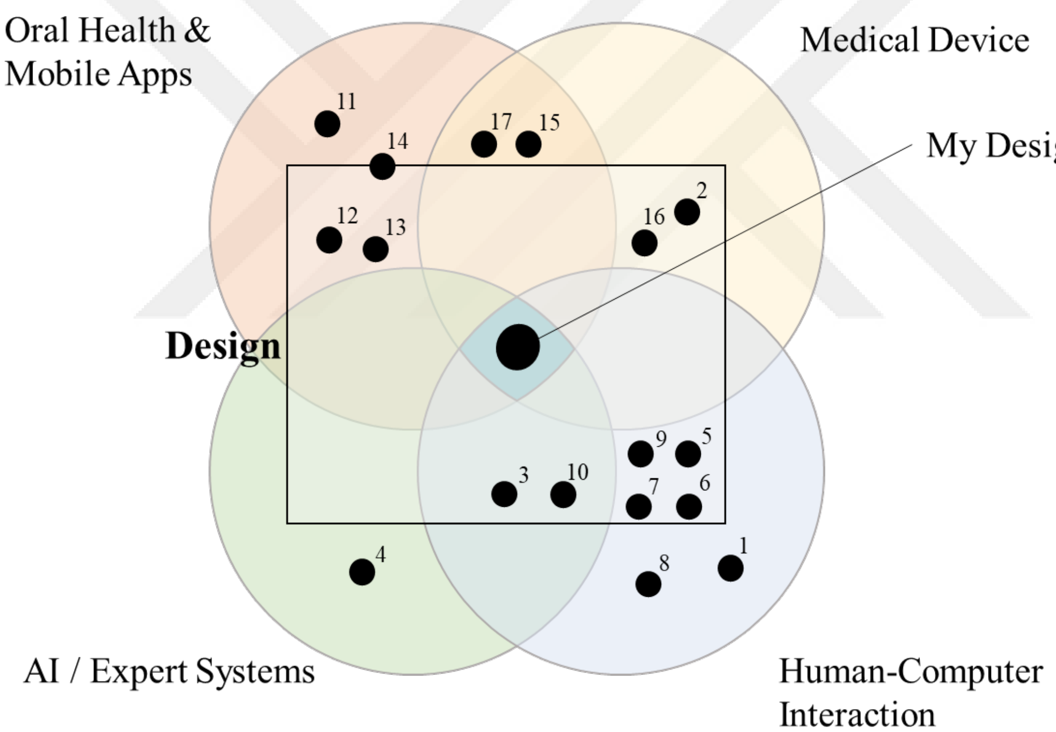
Figure 1.1: The writing styles of the sections.

Figure 1.2 illustrates the map of the four areas related to the thesis and the positions of some sections in the first chapter on this map.

1.2.1 Research Questions

The thesis aims to document the design and development process of an AI-powered diagnostic mobile app for oral health as transparent as possible by looking at it from a design perspective, without losing touch with where I stand as a researcher. I was not an observer, on the contrary, I was an actor, a participant. I was not only a design researcher but also a health expert and an entrepreneur. Therefore, it adopts a hybrid research approach to overcome these conflicts triggered by the complex nature of the project and my multiple identities, just to tell it as it is without pretending.

Since this study was not intended to generate scientific knowledge and did not start with a hypothesis, there is no point in asking questions. Instead, it would be more appropriate to state what it deals with. It deals with an innovation project where social studies and natural sciences, design and dentistry, qualitative methods and engineering, and design research and mDentistry meet. It deals with what I learned from the experience and the literature review and what I reflected.



1- Defining users (the intro)
2- Users in the medical device context
3- Redefining the user in the context of expert systems
4- Introduction to AI and expert systems
5- Designing the interaction (the intro)
6- The role of semiotics in interaction
7- Brief introduction to the Theory of Affordances
8- Introduction to user modeling
9- User profiling
10- A model proposition for profiling the users of expert systems
11- Clinical approach in mDentistry
12- Design approach in the evaluations by health experts
13- The design methods adopted in the development processes of oral health apps
14- Issues of approaches in mDentistry literature
15- Governing regulations for medical apps
16- The role of designers in the medical device development processes
17- Prospects and challenges of regulations

Figure 1.2: The map of the four areas related to the thesis and the positions of the sections on this map.

2. LITERATURE REVIEW

2.1 Human-Computer Interaction

This section consists of three subsections. In the first part, the human being, that is, the user, will be defined, in the second part, a brief introduction to artificial intelligence and its branch, expert systems, will be made. In the third part, the factors affecting the interaction between humans and computers will be discussed and a new model for user profile will be proposed.

2.1.1 Defining users

It would be appropriate to start by defining the human, that is, the user, in accordance with the name of the human-computer interaction section. In this section, I will concentrate on the current user definitions in the HCI literature and redefine the user according to various characteristics in the context of expert systems.

In the early days of computer science research, the focus was more on machine-related errors, while human-factor-related problems were only superficially addressed (Schneiderman, 1976). The need for a separation between these two issues was understood and some researchers focused more on human users to define and elaborate their characteristics. Miller (1974) indicated the influence of prior experience with procedures on the production of program errors. Similarly, Schneiderman (1976) also addressed the crucial role of users' previous experiences in performance and categorized users into four divisions according to the number of programming courses they had taken: "*naïves*" are those who had not attended any programming courses; "*novices*" are those who were currently enrolled in and completed their first programming course; "*intermediates*" are those who were currently taking or had just completed their second or third course; and "*advanced users*" are those who were graduate students and faculty members of a computer science department.

After computers began to be widely used, users who do not have expertise in computer science but who nevertheless can benefit from the use of computers in their work or daily lives became increasingly heterogeneous. At this point, it was needed to distinguish between a user who has just started taking computer courses at university and a user who encounters it at the bank, at work, at school, or home but has no experience. Eason (1976), in his study on managers' interaction with computers, defined "*naive users*" as those who are not computer experts but make direct use of computer systems in the everyday performance of their duties. Coombs and Alty (1980) investigated independent variables affecting the course of interaction in advisory services held in computer centers and found that beginners whom they defined as "*inexpert users*" have greater difficulty in benefiting from the help given by advisors. Scapin (1981) compared two groups of subjects having different levels of experience with computers in terms of learning and recall of computer commands and found that the "*experienced users*" were able to recognize commands, but the "*naïve/inexperienced users*" were not. On the other hand, in their review, Paxton and Turner (1984) made no distinction between the two terms, "*novice*" and "*naïve*" and used them interchangeably. Like Scapin (1981), Burgess and Swigger (1986) have also defined a naïve user as someone who has little or no expertise in operating computers.

Allwood (1986) emphasized that novice users were not uniformly used in the literature; and noted that such users were not the least experienced users in a research group, instead, it would be more accurate to use the term to represent users with little or no experience in computer tasks. He also preferred to call users with experience "*experts*" instead of Schneiderman's term, "*advanced users*". He attributed the reason for this change to the terminology used in studies conducted in other fields (Larkin, 1981, 1983; Simon & Simon, 1978) that showed experts' use of knowledge and how it played an important role in enabling them quickly to tackle the problem they encountered.

Later, Fisher (1991) wanted to make a distinction between the novice and naïve and between experienced and expert terms to put an end to this confusion in the literature. He brought a different perspective on this issue by stating "Thus while the terms, novice and naïve may coincide, they carry separable meanings..." (Fisher, 1991, p. 439) and defined "*naive users*" as those who lack the ability or capacity to analyze and reason in a given situation and who need to make meaningful use of computers for certain tasks but are not interested in other functions; and "*novice users*" as those who are new or inexperienced in a certain task or situation, in other words, "*beginners*" or "*new users.*" By underlining that being experienced does not say anything about the quality of the experience gained, he defined the "*experienced user*" as someone who has developed skills or knowledge due to participation or exposure. According to Fisher (1991), an experienced user is the opposite of a novice user. Lastly, he defined "*expert users*" as those who make an effort to gain skills and knowledge and benefit from them to use the system extensively. In Fisher's terms, being an expert stands for the opposite of being naïve. He stated that even a new user who intends to become an expert would begin as an expert novice[1].

The efforts in the literature to define the user, sometimes using the same and sometimes different terms, have progressed in parallel with the studies revealing the factors affecting the interaction of the user with the computer, and with the fact that computer use has become popular and consequently, the diversity of the user profile increased. Since the first studies were carried out with computer science students, the definitions were determined based on the experience and the number of courses taken by the students. After a while, as computers started to be used by users who did not have the expertise, confusion started in the definitions of the terms. Of course, the necessity of distinguishing between two novice users emerged when one of them was a computer science student, and the other was not. However, the literature could not reach a consensus on this issue either. As studies have been carried out, it has been seen that making definitions is not that simple. Another issue is that a person's experience with computer systems also depends on that person's access to that technology. Therefore, not all potential users can benefit from these technologies equally. Inequality exists not only between developed and developing countries but also between those who have easy access to information and those who do not know how and where to find it (Goulding, 2001). In this regard, Van Deursen and Van Dijk (2009) underscored that users having access to or owning a computer does not necessarily mean that they have the capacity to operate and use it and they define this

[1] Fisher wrote the term in his original paper as "*experienced novice*". Throughout his article, he distinguished between novice and naive, depending on how prone the user is to develop new skills. So, assuming it's a typo, I change it to expert-novice.

divide as "*digital skills*". In fact, this is the subject that Fisher (1991) wanted to emphasize in his work by making the distinction between novice vs. naïve and experienced vs. expert. Users may be inexperienced; their experience may increase with computer exposure, but it is not the only parameter. Users' willingness to use the computer extensively also became prominent.

2.1.1.1 Users in the medical device context

Like computers, as medical devices for the use of patients and their relatives at home became widespread, similar discussions have emerged about defining users. While some terms from the HCI were transferred to the field, new terms were also coined.

Lay-user is frequently used in the medical device literature. Hogg and Williamson (2001) argued that its description is not clear although it is largely used in the health field. In dictionaries, a layperson is defined as *"a person who is not trained, qualified, or experienced in a particular subject or activity*" (Collin's dictionary[2]) and as "*a person without professional or specialized knowledge in a particular subject*" (Oxford Dictionaries Online[3] as cited in Cifter, 2011). In this sense, the definition is made not according to the positive qualities that these people may have, but according to who they are not and what they do not have (Hogg & Williamson, 2001).

In his thesis, Cifter (2011) criticized the implication of specialized knowledge in Oxford's definition by stating that "…in the medical field, patients can develop their own specialized knowledge of their specific condition or their illness through personal experience, or via other information sources such as the Internet, magazines or books*"* (p. 22). Similarly, Entwistle et al. (1998) implied the same issue and stated "…We use the term "*lay*" to mean people who are neither health care professionals nor health services researchers, but who may have specialized knowledge related to health. This includes patients, the general public, and consumer advocates" (p. 463). Cifter (2011) supported Entwistle's argument and noted that although his definition emphasized the knowledge lay user might also have, it was not clear what it meant

---

[2] https://www.collinsdictionary.com/dictionary/english/lay-person (Last accessed: 01/04/2023)

[3] http://oxforddictionaries.com/view/entry/m_en_us1262529?rskey=yqz7F0&result=2#m_en_us 1262529 (last accessed: 18/06/10)

by "*professional*". According to Hogg and Williamson (2001), the term professional covers a wide range of attributes and the term lay covers even wider.

In Cifter's terms (2011), a lay user is "*A user of a product or a system who has not undergone extensive training in the subject field (which enables him/her to be eligible to act as a member of a profession), but uses the system or the product due to his/her special interest or needs*" (p. 24).

A professional user was also defined by highlighting the knowledge and existence of training (Cifter & Dong, 2009). According to them, professional users are *"the users who have gone through extensive training to achieve particular knowledge which is valuable in a social or economical context"* (p. 4).

Cifter, (2011) classified users according to the definitions above and according to three determinants: Training, knowledge, and experience (Figure 2.1). In this classification, the end-users of a medical device were first divided into professional and lay. It was noted that if the lay user is experienced in using the medical device, they may have limited knowledge compared to professional users. Hogg and Williamson (2001) stated that as laypeople become more knowledgeable about the profession and develop a greater understanding, they lose their amateur status, resulting in the loss of representation of other laypersons who do not have the same knowledge.

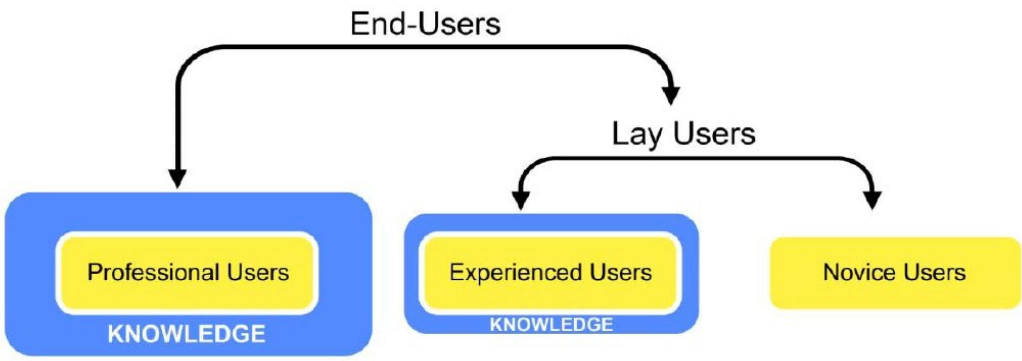


Figure 2.1: The classification of end users (Cifter, 2011, p. 25).

2.1.1.2 Redefining the user in the context of expert systems

Since I thought that the definitions made in both HCI and medical device literature were incompatible with my case (the expert system in pediatric oral health that I developed in the project), I decided to redefine the users in the context of expert systems. All definitions of the user that you will encounter in the following sections of the thesis are according to the definitions I made in this section.

Before defining the users, it would be more appropriate to define the expert system. They can briefly be defined as algorithms that mimic the decision-making process or thinking of an expert from a specific field. Some expert systems are based on scientific literature, while others also draw on the expert's professional experience or both (Gupta & Nagpal, 2020). Since it is software mimicking human behavior, it is accepted as a branch of artificial intelligence in computer science (Abu-Nasser, 2017).

Expert systems, as the name implies, are usually designed to be used by experts. In medicine, they are called "*clinical decision support systems*" and their use purpose is to help or assist practitioners in decision making such as diagnosing, individualized dosing, evaluating laboratory test results, etc. (Wasylewicz & Scheepers-Hoeks, 2019). In this context, we can simply define an expert user, utilizing Cifter's definition of the lay user (2011, p.24), as a person who has the knowledge or expertise from which the system is derived and has undergone training in the relevant field which enables him/her to act as a member of the profession. Owing to that, being an expert user is highly dependent on the area of the system's expertise. Like expert users, non-expert users have the same dependency.

According to the Cambridge Dictionary[4], a non-expert is "*a person who does not have a high level of knowledge or skill relating to a particular subject or activity.*" As Entwistle et al. (1998, p. 463) and Cifter (2011, p. 22) implied above, even if non-experts are not trained in the relevant field, this does not mean that they do not have the knowledge. As in Cambridge's definition, their level of knowledge is just not as high as experts who have been educated in the relevant field. Therefore, we can describe them as users who have not undergone any training and professional

[4] https://dictionary.cambridge.org/dictionary/english/non-expert (Last accessed: 09/04/2023)

experience in the field from which the expert system they use is derived. Although the lay user, in Cifter's terms (2011, p. 24), has a similar meaning to the non-expert user, in the context of expert systems, I believe that defining users through this contrast or antilogy emphasizes the differentiating factor, which is the expertise in the relevant field. For this reason, I prefer to use the non-expert user term instead of lay user on purpose.

As mentioned earlier, novice user is a widely used term in HCI to indicate users with no previous experience and knowledge. However, I believe that this label does not fully represent the non-expert users using expert systems. Because being a novice is solely dependent on experience in using the product regardless of the expertise and tends to change over a period of time as the user continues to use the product. Expert users have training in the relevant field, from which the expert system's knowledge base is derived, so they are familiar with the content. However, being an expert is not always equal to being advanced in using an expert system. Therefore, expert users can also be novices as well as non-experts. It is worthy of note that they are two separate labels that identify two different attributes of the user, and they are not mutually exclusive. For example, being a dentist does not necessarily mean being an advanced user of an expert system in dentistry. They may become advanced faster than non-expert users, but they still start using it as a novice if they have no previous experience with such systems. The same applies to a user who studied computer science. Although she/he may be an expert in computers and software, she/he still uses a medical expert system as a non-expert user. At this point, my classification differs from Cifter's (2011).

As Fisher (1991) underscored, some users' understanding, analyzing, and problem-solving skills can be better than others. Moreover, any impairment in their visual perception, memory, information processing skills (Kaye & Crowley, 2000), or even personality traits such as attention deficit hyperactivity disorder (McKnight, 2010) can negatively affect their ability to use a system. I do not get into the details because the cognitive aspects of skills fall within the scope of psychology and are beyond the scope of this thesis. All I want to emphasize here is the "*internal factors*" (Osvalder & Ulfvengren, 2009) that users inherently possess and somehow affect the interaction. That is what I call the user's competence and I define being "naïve" or “adept” through this axis. According to this classification, both expert and non-expert users can be naïve or adept at using expert systems.

In my classification, I define the user’s competence without including the user’s willingness or openness for extensive use unlike Fisher (1991). Because I believe that they are two separate attributes but not completely unrelated. While one of them has to do with the physical and psychological state of the user, the other is related to the state of being interested in the subject. Just because someone can do something does not mean they want to do it or the other way around. However, some skills can also be improved, but they cannot exceed the individual's maximum capacity.

In summary, I classify end users of an expert system based on three main attributes: expertise, experience, and competence. According to these classes, a user can be either an expert or non-expert depending on the training in the relevant field; a novice or advanced depending on the experience; a naïve or adept depending on the skills (or other internal factors). Table 2.1 illustrates the classification.

The types of users can briefly be described as follows:

Expert Users: They are people who have the knowledge or expertise from which the system is derived and have received training in the relevant field.

Non-expert Users: They are people who do not have any training and professional experience in the field from which the expert system they use is derived.

Novice Users: They are the users who have never used the expert system before and have no experience.

Advanced Users: They are the users who are experienced in using an expert system or accustomed to similar software.

Naïve Users: They are the users who have limited abilities such as reasoning, problem-solving, seeing, reading, etc., that enable users to interact with a product.

Adept Users: They are the users who do not experience any difficulties in learning how to use a new product and who have the required competence to make extensive use of the product's features.

Table 2.1: Classification of end-users of an expert system.

	End-users of an Expert System			
	Expert User		Non-expert User	
Training in the relevant field	Yes		No	
Knowledge	Has expertise in the field		None or to a certain extent	
Experience	Novice	Advanced	Novice	Advanced
Competence	Naive	Adept	Naive	Adept

In the interaction section, I will elaborate on these three attributes with a three-dimensional model in the context of an expert system.

2.1.2 Introduction to artificial intelligence and expert systems

Intelligence is the ability to acquire knowledge with experience and skills to achieve certain goals. The level of intelligence differs in different species such as humans and animals. From an evolutionary point of view, the ultimate goal of living beings is to survive in an environment and adapt their behaviors to ever-changing conditions, so we observe our surroundings and kins to learn and predict the possible consequences of certain actions. According to Sternberg (1998, 2008), intelligence performs best, when it is “augmented” with wisdom. Therefore, intelligent beings must draw inferences from their own or others' experiences and keep them in their memory consistently to use them later to avoid negative consequences in the future or to achieve certain goals.

A simple computer lacks reasoning and common sense because it does not have consciousness and it is not possible for a computer to understand and adapt to new situations unless it is programmed to do so. They are machines that cannot go beyond the functions and physical characteristics determined by the engineers and designers who created them. Humans are conscious beings. They see, hear, smell, taste, and touch. Thanks to these senses, they recognize the objects and living things in the environment and can interact with others. They gain experience and knowledge through their observations and actions, and they store the information they learn. At this point, artificial intelligence (AI) aims to bring all these human features to computers to perform more complex tasks.

The term, artificial intelligence, was first coined by McCarthy (2006), a computer and cognitive scientist, at the Dartmouth Conference in 1955. He and his colleagues suggested that a machine could mimic the ability to learn and any other aspects of intelligence that were then reserved only for humans. Later, other definitions were proposed such as "*The art of creating machines that perform functions that require intelligence when performed by people*" by Kurzweil (1992, as cited in Gupta & Nagpal, 2020, p. 10) and "*The branch of computer science that is concerned with the automation of intelligent behavior*" by Luger and Stubblefield (1997, as cited in Gupta & Nagpal, 2020, p. 10). It is simply a man-made system that is expected to behave like humans, capable of perceiving and recognizing (speech and image recognition), processing, acting, and reacting. Some intelligent systems are designed for problem-solving, planning, and decision-making, and some are capable of self-learning and -updating, which is called machine learning.

The test to evaluate how successfully artificial intelligence can mimic a human was first devised by Alan Turing (1950). The Turing Test named after him, aims to determine whether an intelligent machine can achieve human-level performance in given tasks to sufficiently fool the interrogator. In the Turing Test, there are three terminals: two humans and a machine in a space where they are separated from each other. One of the human terminals, the interrogator, is designated to ask a series of questions and the other two terminals, the human, and the machine respond. After the interrogation, if the interrogator cannot distinguish which answers are given by the machine, the intelligent system passes the test.

However, conducting such a test is not as easy as it seems. First, the results largely depend on the type of message. The more diverse messages the interrogator receives such as the audio or visual, the harder it becomes to pass the test. Second, the type of questions is another parameter that affects the results. It would be difficult for the interrogator to identify the machine from the responses to yes/no questions. Another issue to consider is how the questions are formulated because humans are much better than machines when it comes to conversations, on the contrary, machines can easily beat humans in mathematical calculations. According to Gupta and Nagpal (2020, p. 7), there are four capabilities that AI must possess to pass the Turing Test:

1) Natural Language Processing to understand what the interrogator says.

2) Storing the information provided before and during the interrogation.

3) Reasoning to answer questions or draw new conclusions leveraging the stored information.

4) Learning to adapt to new circumstances by detecting and extrapolating patterns.

Expert systems are programs that are studied in computer science under the roof of the research branch of AI. An expert system is simply an algorithm that mimics an expert's decision-making ability in a particular field, in other words, it is "*an application of AI that embodies human expertise*" (Gupta and Nagpal, 2020, p. 71). As Sternberg (1998, 2008) earlier pointed out, intelligence shows high performance with knowledge. Expert systems are also known as knowledge-based programs. In the design process, a knowledge engineer who studies how experts make decisions translates what is known in their expertise into a rule-based system, so that a computer can simulate their decision-making behavior.

An expert system contains five main components (Figure 2.2):

1- Knowledge base

2- Database or storage

3- Inference engine

4- Explanation facility

5- User interface

2.1.2.1 Knowledge base

The knowledge base is the representation of the knowledge in a particular domain in the system, which contains information about relevant parameters, the rules that describe the relationship between these parameters, and the outcomes. It plays a key role in determining which outputs will be shown under which conditions or after which inputs, the user enters into the system.

According to Gupta and Nagpal (2020), three types of knowledge can be applied in an expert system: Procedural, factual, and heuristic. Procedural knowledge-based systems perform all kinds of tasks step by step related to the main task that is aimed to be completed at the end of the process. For example, a robot designed to open a locked door must first pick the right key, insert it into the lock, and turn it. If everything goes well with the sub-tasks, the system manages to complete the main task, which is to open a locked door. The factual knowledge-based systems draw upon the facts of a particular domain that can be found in science textbooks or journals. Expert systems in medicine often use factual knowledge. For this reason, decisions that are critical to human life cannot be left to systems developed with subjective rules, because their harmful consequences can be enormous. The heuristic systems on the other hand are experimental and subjective.

The rules in the knowledge base consist of two major parts: conditions (antecedents) and conclusions (consequents) in the form of "IF" and "THEN". The conditions and conclusions respectively come after "IF" and "THEN". To avoid conflicting outputs, the knowledge engineer should first understand the expert's way of thinking, what conditions are necessary to draw a particular conclusion, and especially the distinctive features of each possible outcome.

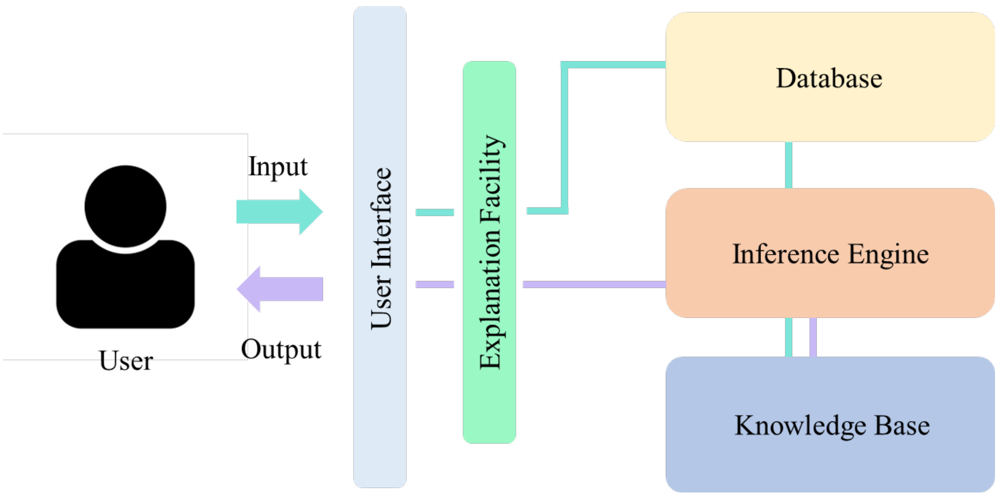


Figure 2.2: The components of an expert system.

For example:

Let's imagine a simple algorithm that suggests what to wear before going out, depending on the weather. The conditions are sunny, cloudy, and rainy, and the consequents are a t-shirt, a jacket, and a raincoat as shown in Figure 2.3. The rules can be as follows:

- IF it is sunny, THEN he/she should wear a T-shirt.
- IF it is cloudy, THEN he/she should wear a jacket.
- IF it is rainy, THEN he/she should wear a raincoat.

The inference above is simple, and the distinctive features of each parameter are very clear. There are three mutually exclusive premises in the algorithm: whether it's sunny, cloudy, or rainy, and depending on each circumstance, there is a related consequence.

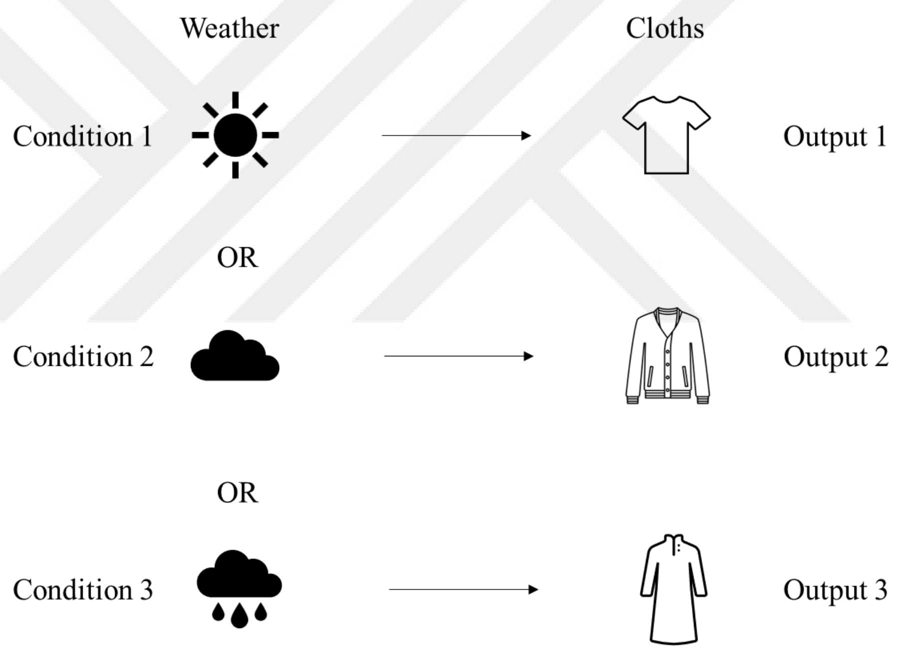


Figure 2.3: The rules of the example algorithm that suggests what to wear depending on the three weather conditions.

Rules can contain multiple antecedents that can be joined with logical connectives such as "AND" (the conjunction connective) or "OR" (the disjunction connective) and become compound statements. For example, when the two antecedents are conjoined with “AND”, the consequent part depends on these two being true at the same time. On the other hand, the consequent in an “IF…THEN” statement depends on either one of the antecedents being true, that is conjoined with “OR”. The relationship between weather conditions in Figure 2.3 is disjunctive, so it is either sunny, cloudy, or rainy.

If we add a second antecedent such as temperature to the rules using the conjunction connective "AND", and diversify the consequents (Figure 2.4), the rules become as follows:

- IF it is sunny AND hotter than 22ºC, THEN he/she should wear a t-shirt.
- IF it is cloudy AND hotter than 22ºC, THEN he/she should wear a t-shirt.
- IF it is rainy AND hotter than 22ºC, THEN he/she should wear a T-shirt and a raincoat.
- IF it is sunny AND colder than 22ºC, THEN he/she should wear a T-shirt and a jacket.
- IF it is cloudy AND colder than 22ºC, THEN he/she should wear a jacket.
- IF it is rainy AND colder than 22ºC, THEN he/she should wear a raincoat.

The first two rules can also be stated as follows:

- IF it is sunny OR cloudy, AND hotter than 22ºC, THEN he/she should wear a t-shirt.

As shown in the rules, there is always an output for every possibility. If the system cannot find the relevant rule in the knowledge base for the valid set of conditions, it cannot infer, resulting in an error.

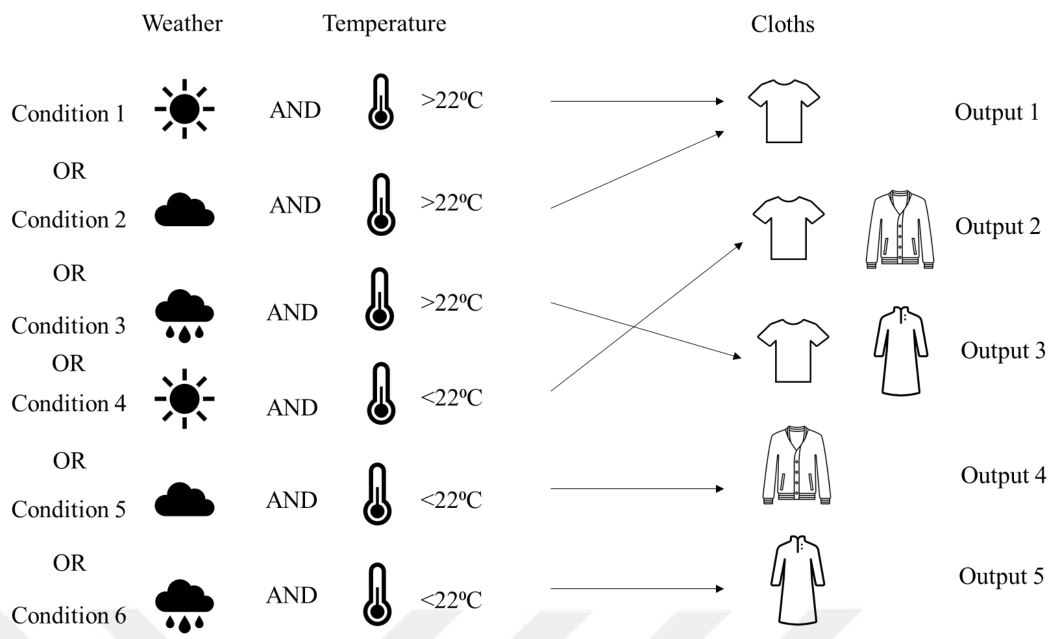


Figure 2.4: The rules of the second algorithm that suggest what to wear depending on the weather conditions and the temperature.

2.1.2.2 Database or storage

A database or storage is a component in an expert system, that acts as a memory, holding user-entered inputs related to the problem being solved, to be processed by the inference engine. The inference engine uses the input from the database, scans the knowledge base to find a matching antecedent part of a rule, and if they match, it draws a conclusion from the existing knowledge. The database stores the relevant inputs, intermediate results, and conclusions. In some cases, it pulls data from external sources. The difference between the database and the knowledge base is that the database holds the data related to the particular problem, whereas the knowledge base contains information from the domain to be used repeatedly for other problems (Gupta and Nagpal, 2020).

2.1.2.3 Inference Engine

The inference engine works as an interpreter that compares the inputs the user enters into the system with the antecedent part of the rules. If there is a match, it infers a conclusion drawing the consequent part of the relevant rule. There are two kinds of inference methods: Forward chaining and backward chaining. In the forward chaining inference method, the inputs are first processed by the rules, while in the backward chaining inference method, the goal or consequent is processed initially, so that the system looks for the relevant rules and necessary antecedents to conclude that determined goal. Returning to the same example, a forward-chaining inference engine would query the weather conditions to recommend which clothes the user should better wear. In the backward chaining inference engine, the user must decide first what to wear, and then the system checks what conditions are necessary to achieve the goal.

Considering that the examples given in the current literature may be complex for readers unfamiliar with the subject, I will simply explain both methods with examples.

The first example:

In the forward-chaining inference method, the question order is as follows:

1- Ask about the weather

2- Ask about the temperature

The user flow is shown in Figure 2.5 and the algorithm works as follows:

STEP-1: Inference Engine → Explanation Facility: Prompt the first question.

STEP-2: Explanation System → User Interface:

How is the weather?

a) Sunny

b) Cloudy

c) Rainy

STEP-3: User's Input: a → User Interface → Database: Sunny is true.

STEP-4: Database → Inference Engine: Sunny is true.

STEP-5: Inference Engine → Knowledge Base: If Sunny is true, then…

- Wear a T-shirt. (AND if it is hotter than 22ºC)

OR

- Wear a T-shirt and a jacket (AND if it is colder than 22ºC)

STEP-6: Inference Engine → Explanation System: Prompt the second question.

STEP-7: Explanation System → User Interface: Is it hotter than 22ºC?

- Yes
- No

STEP-8: User → Input: Yes → User Interface→ Database: Hot is true.

STEP-9: Database → Inference Engine: Hot is true.

STEP-10: Inference Engine → Knowledge Base: If Sunny AND Hot are true, then…

- Wear a T-shirt.

STEP-11: Inference Engine → Explanation System → User Interface: Wear a t-shirt.

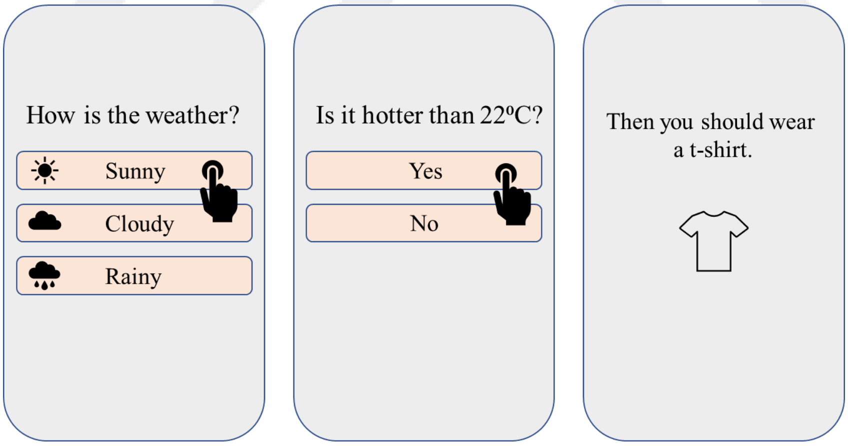


Figure 2.5: The user flow of a system using the forward chaining inference engine.

The second example:

In the system using the backward chaining inference engine, the question order is as follows:

1- Ask him what he would like to wear today.

2- Ask about the temperature.

3- Ask about the weather.

The user flow is shown in Figure 2.6 and the algorithm works as follows:

STEP-1: Inference Engine → Explanation System: Prompt the first question.

STEP-2: Explanation System → User Interface:

What do you want to wear today?

a) T-shirt
b) Jacket
c) Raincoat
d) T-shirt & Jacket
e) T-shirt & Raincoat

STEP-3: User’s Input: a → User Interface → Database: T-shirt is true.

STEP-4: Database → Inference Engine: T-shirt is true.

STEP-5: Inference Engine → Knowledge Base: Then the T-shirt is true, if…

- It is hotter than 22ºC.
- It is colder than 22ºC.

STEP-6: Inference Engine → Explanation System: Prompt the second question.

STEP-7: Explanation System → User Interface: Is it hotter than 22ºC?

- Yes
- No

STEP-8: User’s Input: Yes → User Interface → Database: Hot is true.

STEP-9: Database → Inference Engine: Hot is true.

STEP-10: Inference Engine → Knowledge Base: Then T-shirt is true, if Hot is true AND if it is…

- Sunny

  OR

- Cloudy

STEP-11: Inference Engine → Explanation System: Prompt the third question.

STEP-12: Explanation System → User Interface:

How is the weather?

a) Sunny

b) Cloudy

c) Rainy

STEP-13: User's Input: a → User Interface → Database: Sunny is true.

STEP-14: Database → Inference Engine: Sunny is true.

STEP-15: Inference Engine → Knowledge Base: Then T-shirt is true if Hot and Sunny are true → The rule exists.

STEP-16: Inference Engine → Explanation System → User Interface: You can wear a t-shirt.

Both methods can be used for different purposes. In the medical field, expert systems with a forward chain inference engine can be used when diagnosing an unknown disease, while those with a backward chaining inference engine can act as diagnostic protocols to confirm the possible diagnosis made by healthcare professionals.

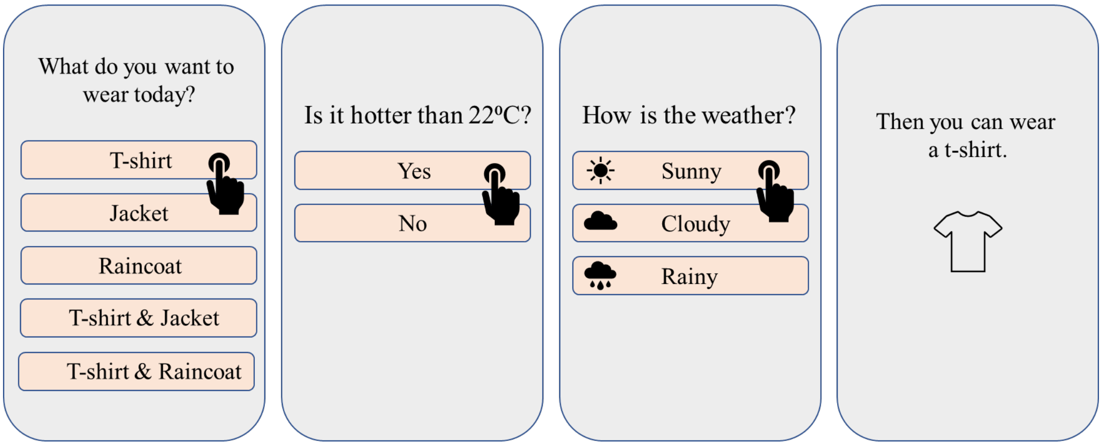


Figure 2.6: The user flow of a system using the backward chaining inference engine.

2.1.2.4 Explanation facility

As noted in the literature (Gupta & Nagpal 2020), the explanation facility is a component that informs the user or developer of how the expert system achieves certain results by using which rules played a role during the inference process. The explanation facility enables users to backtrack on which data and rules the system has used in the process to get the output, so it increases the system's accountability (Gaines, 1991).

However, an additional function can be defined for this component. The knowledge base contains the outputs to be obtained as a result of the decision process and the rules that determine which premises these outputs depend on. So, in which part of this system do the questions that enable these inputs to be drawn from the user, and the possible answers to these questions take place? Whether the user is an expert or a non-expert and whether she/he understands what the inputs mean plays a key role at this point. If an expert system is designed for use by non-experts, there should be an intermediary component that allows the user to enter the correct input into the algorithm by explaining which input corresponds to what to increase the accuracy of the algorithm. Thus, the explanation facility can take over this function.

2.1.2.5 User interface

The user interface is the component that allows the system to communicate with the user, it is where the interaction takes place. They give orders to the system through interfaces or answer the questions asked by the system. Users are generally not concerned with what is inside the program and how it works, but with how it communicates with them. Communication has several dimensions including content, information architecture, navigation, graphics, layout, and the design of interface signs such as labels, affordances, and icons (Bolchini et al, 2009). Therefore, interface design in all its scope plays a critical role in the system's usability.

Expert systems have several advantages and disadvantages when compared to other artificial intelligence methods. The advantages of expert systems:

- Reliability: Expert systems are highly reliable systems if the knowledge base is developed without bias and subjectivity (Gupta & Nagpal 2020).

- Efficiency: The number of rules or the data volume is not a major concern for the computational resources used by the algorithm (Gupta & Nagpal 2020).

- Transparent: It can be easily determined in which situations the system reaches a certain conclusion. The rules, inputs, and outputs are clearly compared to other complex algorithms (Gupta & Nagpal 2020).

- Updateability: Developers can change or delete the rules and add new data to the knowledge base over time.

- Practicality: Well-designed expert systems are practical and useful. They can act as reminders of forgotten or underrated facts in a field.

The disadvantages of expert systems:

- Incapable of Learning: Expert system algorithms are incapable of learning new subjects. They have no common sense or reasoning, so they cannot create new rules on their own, which makes them dependent on the developer.

- Lack of Creativity: Expert systems cannot adapt themselves to new conditions or respond to inputs differently from the way they were formulated. They strictly adhere to the rules (Gupta & Nagpal 2020).

- Limited Inference: Expert systems only infer the results formulated in the knowledge base. The parameters, weight points, calculations, and results are predetermined until the developer makes changes. In this sense, their maintenance is expensive (Gupta & Nagpal 2020).

2.1.3 Designing interaction

The interface is a medium where the user interacts with the system or the machine. Although every product has an interface, when it is mentioned, it usually comes to mind as an element of software products. In software, interface design is critical in terms of interaction and usability. HCI is a multidisciplinary field that concerns and studies the factors affecting the human use of interactive computing systems (ACM SIGCHI, 1992) in which computer science, cognitive psychology, human factors, engineering, design, and social and organizational psychology contribute (Liu et al, 2010). For the past decades, a wide range of studies have been conducted on various aspects of HCI such as dialogue techniques, gestural analysis, multimodal interfaces, computer graphics, computational linguistics, spatial cognition, robot navigation and wayfinding, input styles or devices, monitor screens, etc. but the ultimate goal of these studies has always been improving the interaction, usability, and user satisfaction (Liu et al, 2010).

During the interaction, various factors arising from the user, the task, the technical system, and the work environment can affect users' overall performances (Liu et al, 2010). Osvalder and Ulfvengren (2009) classify such factors into three categories: internal factors that are conditions affecting human operators mentally or physically such as age, perceptive and cognitive abilities, personality, and motivation, etc., external factors that refer to operational and environmental issues within the working contexts such as conditions of surroundings, tools and equipment, interfaces, etc., and lastly, stressors that negatively affect the operator's decision-making process directly or indirectly, such as high workload and pace, pain, exhaustion, etc. Design mostly focuses on external factors, interfaces in particular. Interface design plays a major role in the system's usability (Islam, 2013). Usability is defined as "*the extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency, and satisfaction in a specified context of user*" (ISO, 1998). According to this definition, if the product's target user can use it effectively and achieve their goals without any problems or setbacks, it can be said that the product is usable.

Bolchini et al. (2009) indicate that there are several dimensions of interface design: "content (what is communicated), information architecture (the structuring of information in a meaningful way), navigation (how the information architecture can be used and interacted with), graphics and layout" (p. 67). They consider signs used on the interface as "communication artifacts" that indicate or convey the presence of content, functions, and links (Bolchini et al, 2009, p. 67). Interfaces mostly consist of signs that come into the body in different forms such as links, images, buttons, icons, or animations, and are assigned various functions (Neumuller, 2001). Islam (2011) emphasizes the importance of signs by stating that if these dimensions are the same in two different interfaces, the one with more intuitive and well-used signs can be assumed to be more useful.

During the interaction, communication takes place between two human proxies of systems; on one side the designer sends messages to subtly show how the system works or how the goals can be achieved, on the other side, the user receives and interprets these messages through these beacons. In the literature, the communication between these two sides is called "*metacommunication*" or "*computer-mediated human communication*" (de Souza, 2005; 2018, p. 45). In this sense, it can be assumed that the friendliness of the design is strongly related to the fact that how well the message is formed through the signs in the interface by the designer, and more critically, whether they are all received and interpreted as intended by the user. The key here is the degree to which the designer speaks the same language as the user. It was noted that semiotics can help designers enhance their power in communication (Barbosa et al, 2001).

2.1.3.1 The role of semiotics in interaction

Signs are vehicles to transmit a message or information from someone who knows it and wants others to know it as well (Eco, 1988). Semiotics is the research field that studies "*signs, signification, and signifying systems*" as well as representation, meaning, and reference (Robert et al, 1992, in cited as Islam, 2013). Eco (1976) states "Semiotics is concerned with everything that can be taken as a sign*"* and defines it as a method of investigating how sign vehicles may function and a means of understanding how a sign vehicle may be produced and interpreted.

There are two major figures in the history of semiotics: Ferdinand de Saussure (1857-1913) and Charles Sanders Peirce (1839-1914). Saussure (1983) developed a model in which a sign is composed of two essences: the "*signifier*", the form the sign takes, and the "*signified*", the concept or meaning the sign represents. He calls the relationship between the two the "*signification.*" For example, when the perceiver encounters a "closed" sign hung on the door of a shop, he/she interprets the sign as the shop is closed for customers. What the perceiver sees on the door is considered a "*signifier*", while what he/she interprets refers to a "*signified*" in Saussure's model. Peirce (1931, p. 58), on the other hand, developed a triadic model that consists of three essences: the "*representamen*", the "*object*" and the "*interpretant.*" Representamen is the representation of an entity, for example, if the entity is a person, it addresses him/her by creating an equivalent or a more advanced version of that person in the perceiver's mind (Ferreira et al, 2005). The object is the actual entity the sign represents (Peirce, 1931). The interpretant is the meaning of the sign created in the perceiver's mind or the reaction caused by the object in the perceiver (Andersen & Nowack, 2002, as cited in Ferreira et al, 2005). For example, the diskette icon or the save icon in the form of a diskette is the representamen which represents the object of saving the document and is interpreted as "*this sign is for saving my document file*" (Islam, 2013, p. 48).

Peirce (1931) classified signs into categories and the most acknowledged and fundamental sign classes are icon, index, and symbol. If the representamen resembles or in some way imitates the object, the sign is considered iconic. Ferreira et al. (2005) exemplify iconic signs with the portrait of Charles Sanders Peirce. When a perceiver sees it, he/she can easily recognize it (if she/he knows his appearance) as his portrait, because the representamen resembles the object (the real Peirce) enough (Ferreira et al, 2005). If there is a causal relationship between the representamen and the object, the sign is indexical. In such cases, the representamen creates a link that refers to the object in the perceiver's mind. For example, a human footprint (representamen) in the snow indicates someone who has walked here (object) rather than the outsole of a shoe. If the relationship between the object and the representamen must be learned before it can be interpreted by the perceiver, the sign is considered symbolic (Ferreira et al, 2005). Symbols can also be defined as objects that create meanings for a particular group of people or culture (Luna et al, 2008).

Before making inferences, one must know in advance that the cross is the symbol of Christianity because nothing in the *representamen* allows the perceiver to interpret the cross in this way. In fact, being identified is a prerequisite for all kinds of signs as Peirce (1931, p. 58) stated "Nothing is a sign unless it is interpreted as a sign." If a human footprint cannot be identified, it would not be possible for the perceiver to infer the footprint is a sign of human presence.

The relationship between the object and the representamen is not like black and white, but rather a spectrum, and the three sign classes are not mutually exclusive, so a sign may contain elements of iconicity, indexicality, and symbolism at the same time to varying degrees (Peirce, 1931). A photograph can be given as a well-known example of a sign carrying all. However, one type of element tends to dominate in the sign, and when it happens, that type becomes prominent (Peirce, 1931). Likewise, the link between the object and the sign is not always monotony; some signs may have a single meaning, and some may have more than one (Islam, 2013). Perceivers often make assumptions about the meaning of the sign according to the context, for example, the C sign can be interpreted as the carbon element in chemistry or a programming language in computer science (Islam, 2013). In this sense, it can be said that meanings are contextual and derived from culture and not always universal, hence, a sign may be interpreted with different meanings by perceivers from different socio-cultural backgrounds (Guillemette & Cossette, 2006).

In the complex world of signs with different representations and meanings, designers create interfaces to convey messages to users through them, to make possible the achievement of human goals in the systems (Nadin, 1988). Signs take different forms in the interface as buttons, words, scroll bars, tabs, images, etc. They all contribute to the communication process, given that designing can be considered "*a semiotic activity*" (Ferreira et al, 2005, p. 3). The purpose of the semiotic approach in design is to enable researchers to investigate the use of language/signs in interfaces and their effects on usability (Islam, 2013). Semiotics can benefit not only in increasing communicability but also in increasing usability and user satisfaction; it can assist designers in designing interfaces and hence, in improving users' task performance (Islam, 2013).

Barr et al. (2002) took the interface elements as signs and analyzed them according to the Peircean model. They noted that representamen corresponds to the forms of the signs, the object corresponds to the underlying functionality, and the interpretant corresponds to what was generated in the mind of the user (Barr et al, 2002). Whether the user can successfully perform the intended task during interaction with the system depends on the congruency between the meaning (interpretant) and the sign's function (the object) because if the user gets the wrong idea, it means that there has been a failure in the sign use and its representamen (Barr et al, 2002).

Icons are commonly used in software systems because people generally recognize pictorial descriptions of the same things more easily than their verbal descriptions, they are more cross-cultural than language and occupy less space in the interface (Apple Computer Inc., 1992). Barr et al. (2002) defined icons in computer interfaces as another type of sign with a specific purpose and evaluated them to elucidate their strengths and weaknesses. They noted that the benefit of iconic and indexical icons is that they have the potential to be perceived and understood in the same fashion by users from all around the world with minimal training because of their resemblance (iconic icons) with the actual objects and their causal link to the functionality (indexical icons) having strong ties with our perception of the physical world (Barr et al, 2002). However, they underscored that the *representamen* of these types of icons is subject to physical constraints, therefore, choosing the right representation can be challenging for designers (Barr et al, 2002). On the other hand, they also noted that symbolic icons are arbitrary in nature, and they do not have any connection with the physical world (Barr et al, 2002). This implies that the nature of this type of icon not only allows designers to design them more easily but also requires users to learn what they mean beforehand and how to use them (Barr et al, 2002).

Ferreira et al. (2005) analyzed the interfaces of three sample software semiotically and redesigned the problematic elements they detected. As a result, they stated that the more symbolic signs are used in the interfaces, this will cause problems in navigation and functionality of the system for novice users because symbolic signs require such users to learn by trial and error how to perform their tasks (Ferreira et al, 2005). They also stated that if almost all the signs in the interface are associated with a function, there is a causal relationship, thus meaning that they can all be considered indexical, but how the function is represented by the sign redefines itself. (Ferreira et al, 2005). For example, in an interface of graphic design software, a magnifying glass icon usually represents the function of zooming in and out. In that sense, the sign of this icon is indexical. However, if this icon resizes the object, it turns into a symbolic sign because the real magnifying glass does not change the size of the object. Misuse of such metaphors leads to misunderstandings. When the behavior implemented by the designer is not the same behavior expected by the user, this causes a usability problem (Ferreira et al, 2005).

2.1.3.2 Brief introduction to the theory of affordances and its effect on interaction

As mentioned earlier, well-established communication is vital for successful interaction. A design needs to be intuitive as well as logical, allowing us to properly understand what we can do with what we see (Kaptelinin, 2014). In this sense, intuitive design is a design that speaks for itself, and this is closely related to the concept of affordance.

Affordance is essentially the perception that an object creates in us about what actions we can perform with it. Its theory was first coined by American psychologist James Jerome Gibson (1904-1979) in the field of ecological psychology. He introduced this new concept by denoting what the environment offers and provides to animals for good or bad (Gibson, 1979). Gibson believes that affordance is determined by the harmony between the physical characteristics/anatomy or abilities of animals and the physical characteristics of the environment. For example, steeps are climbable for goats but not for fish, while underwater is livable for fish, but not for goats. That is, mountains afford goats and seas afford fish or vice versa. The idea of actor-environment mutuality makes an inseparable pair. We cannot imagine fish without water. Herein, affordances are objective because they do not depend on value, meaning, or interpretation, yet are subjective as well due to the need for an actor as a frame of reference (McGrenere & Ho, 2000).

Humans, on the other hand, have altered the shapes and materials of the environment to make more of what benefits them and less of what harms them to make life easier for them and their offspring (Gibson, 1979). According to Gibson, this newly created environment is not artificial, it is the same old environment modified by humans (Gibson, 1979). The hands play an important role in modifying the environment to make it affordable for human beings because they provide a wide variety of behaviors. Gibson (1979) highlights the relationship between the anatomy of the hand and the affordance of objects as follows:

> …Objects can be manufactured and manipulated. Some are portable in that they afford lifting and carrying, while others are not. Some are graspable and others not. To be graspable, an object must have opposite surfaces separated by a distance less than the span of the hand. A five-inch cube can be grasped, but a ten-inch cube cannot. A large object needs a "handle" to afford grasping… (p. 57).

According to Gibson (1979), affordance is not something that changes over time as the observer's need changes and its existence is binary; either it exists, or it does not. Independent of being perceived, affordance is invariant and always present there to be perceived (Gibson, 1979). Torenvliet (2003, p. 14) gives a good example of it: "…Even in a dark room where I could not see my chair, it would still be sit-able for me…" So, how do people know that a chair affords sitting? Do they learn affordance? The notion of learning was not comprehensively mentioned in Gibson's theory, rather, the issue of how people learn to perceive affordance was avoided (Kaptelinin, 2014). The existence of affordance and the information that specifies it are two separate things. Its existence is independent of the actor's experience and culture, whereas the ability to perceive it may be dependent on these factors (Gaver, 1991; McGrenere & Ho, 2000). To Gibson, the information is there to be "picked up" by the actor, the only thing is whether he/she is capable of doing it. For example, a secret door in a room still affords going through even though there is no information showing the possibility of that action (McGrenere & Ho, 2000). Later, Gibson's wife Eleanor Gibson and Anne Pick (Gibson & Pick, 2003) asserted that affordances do not automatically present themselves to the actor, therefore, they must be explored through perception and learned with patience and time.

The theory of affordance was later adopted by the design world and became popular in HCI. Norman (1988) first introduced the concept to the design community. He defined it as "*strong visual clues to the operation of things*" and something that can be taken advantage of by designers to enable users to know what to do just by looking without any instruction (Norman, 1988). Affordances show the possibilities of action in the use of a product and are what designers must take into account (Norman, 1988).

Unlike Gibson, Norman divides affordance into two properties of the object: the perceived and the actual, implying that what is perceived is not necessarily the actual (Norman, 1988, p. 9). He states that affordances result from our interpretations, past knowledge, and experience (Norman, 1988). For Gibson, the frame of reference is the action capabilities of the actor, whereas, for Norman, it is the actor's mental and perceptual capabilities, and to him, the affordance would only exist if the information specifying the possibility of the action is perceived and if the actor knows how to use that information (McGrenere & Ho, 2000).

Some authors stated that Norman's definition was imprecise and lack of details (Kaptelinin, 2014), and his deviation from the Gibsonian definition caused confusion and misuse in the HCI community (McGrenere & Ho, 2000). Norman, himself, said that it was his fault and clarified that when he talked about affordance, he meant the perceived affordance which "The Psychology of Everyday Things" (Norman, 1988) is all about (Norman, 1999). McGrenere and Ho (2000) attributed this difference to the fact that Gibson and Norman had related but different goals. According to them, while Gibson focused on humans and animals modifying their environment to change what it affords them, Norman was particularly concerned with how this modification took place and the manner of making the utility more easily perceived (McGrenere & Ho, 2000). Because to him, what actions are possible for the user was more valuable for designers than what actually exists (Norman, 1999).

As mentioned earlier, Gibson's affordance is binary; a stair is either climbable or not and there is no gray area for being easy or hard. He does not address a range, because they are both climbable at the end. Whereas the concept of easiness is what design is concerned with regarding usability. At this point, it is critical to differentiate the two aspects of design: designing the affordance and designing the information that specifies the affordance, in other words, designing the utility/usefulness and designing the usability (Landauer, 1995, as cited in McGrenere & Ho, 2000). The usefulness of a design is determined by the functions the design affords and whether these affordances allow users to perform the necessary tasks and accomplish their goals (McGrenere & Ho, 2000). The usability of a design is determined by how the information that specifies these affordances is conveyed (McGrenere & Ho, 2000). Both are the designer's job (Norman, 1998).

Gaver (1991) also attempted to bring Gibson's idea into the HCI world. He coined several new terms for affordance such as "*sequential*" or "*nested*" and divided it into four categories based on the availability of information for perception (Gaver, 1991). In his paper, he gives several examples of affordances in interfaces. For instance, onscreen buttons appear to protrude from the underlying surface, which gives it a three-dimensional look and depth. The protuberance of the buttons affords press-ability, so users can intuitively understand that they can click on it by using a mouse. Another example is the affordance in scroll bars. A scroll box, the portion indicating the location of the user in a window, differs from the scroll bar it is in, with its different color and the vertical shaft enclosing the bar affords one sort of movement which is up or down. According to Gaver, the look of the bar affords dragging which leads to another affordance, uncovering the rest of the window. He calls such affordance that can be revealed over time "*sequential*" (Gaver, 1991).

In software, the functions such as clicking, scrolling, writing, spinning, dragging, dropping, etc., become invokable by the affordances (McGrenere & Ho, 2000). The raised look of a button affords clicking however, users do not click it just because there is a button. They are interested in the function that will be invoked when the button is clicked on, and they usually understand what the function is from the signs on it (McGrenere & Ho, 2000).

2.1.3.3 Introduction to user modeling

Early on, HCI focused on interfaces, particularly semiotic elements, to find a way to design more usable systems, and as HCI matured, its concerns started to shift beyond the interface (Fischer, 2001). As it became better understood that the factors affecting the interaction were not only dependent on the system but also some human parameters, in other saying, "*internal factors*" (Osvalder & Ulfvengren, 2009), some studies have begun to attempt modeling users of systems. Such studies on user modeling were concerned with tasks, the sequence of operations, and the stages of the thought processes of the users while performing the tasks to explain and predict human behaviors. Although user models are beyond the scope of this thesis as they are the subject of cognitive psychology in HCI, they will be summarized in this section as an introduction.

During the Second World War, it became very important to optimize the operator's performance in the use of newly developed military systems (Biswas & Springett, 2018). Therefore, since then, researchers have tried to simulate the operators to explore their limitations. The concept of modeling the users' thoughts came to the forefront and Newell and Simon (1961) attempted to decompose human problem-solving activity with a computer program called "*General Problem Solver*" (GPS) to simulate the process humans go through to accomplish tasks. As computer systems became accessible to more people, HCI gained more attention in the late '70s (Biswas et al, 2012). Card, Moran, and Newell (1983) were concerned with how humans interact with computers and thought that cognitive psychology was sufficiently advanced to play an important role in the field. They proposed a notion called the "*Model Human Processor*" in which they divided the human processor into three subsystems: the cognitive, the perceptual, and the motor (Card et al, 1983, p.24). The model human processor is one of the earliest models applied in HCI, and still, the most popular one, which aims to decompose human information processing (Biswas & Springett, 2018). There are two practical versions: the "*Goals, Operators, Methods, and Selection rules*" (GOMS) (Card et al, 1983, p. 139) and "*The Keystroke Level Model*" (KLM) (Card et al, 1983, p. 259).

The GOMS model is based on the fundamental principle of task analysis, which is called "*The Rationality Principle*" (Card et al, 1983, p. 86). A person simply does the tasks that are required to achieve his/her goals and the complexity of human behavior does not derive from the complexity of humans most of the time, rather, it derives from the complexity of the task environment (Newell & Simon, 1972). Therefore, one should analyze a task to understand and predict the course of human behavior (Card et. al, 1983, p.86). The rationality principle aims to structure such behavior with a set of formulations: "Goals + Task + Operators + Inputs + Knowledge + Process-limits → Behavior" (Card et al, 1983, p. 86). While the "Goals + Task + Operators" part takes into account only the objective situations; the rest corresponds to hidden constraints such as perception, knowledge, and computational skills (Card et al, 1983, p. 86). The GOMS model decomposes the user's cognitive structure into four components: (1) a set of Goals, (2) a set of Operators, (3) a set of Methods, and (4) a set of Selection rules (Card et al, 1983, p. 140). As an example, they use the model to analyze user behavior in a manuscript editing program (Card et al, 1983, p. 140).

KLM (Card et al, 1983) is a simple model that allows researchers to decompose the interaction with six primitive operators: pressing a key; moving the pointing device to a specific location; making pointer drag movements; performing mental preparation; moving hands to appropriate locations; and waiting for the computer to execute a command (Biswas & Springett, 2018). The duration of these six operations was empirically determined and with the help of this model, researchers can predict the task completion time by the number of times each type of operation must occur to accomplish the task (Biswas & Springett, 2018). This model can be used to specify the user's interaction details.

Other models were brought forward such as Natural GOMS Language (NGOMSL) by Kieras (1994), Cognitive Complexity Theory (CCT) to model the knowledge of users of an interactive computer system (Bovair et al, 1990), GOMS Language Evaluation and Analysis (GLEAN) to simulate the interaction between a simulated user and a simulated device (Kieras et al, 1995), CPM GOMS to explore parallelism in users' actions (John & Kieras, 1996) and so on (for further information look at Biswas & Springett, 2018).

Like in cases where the signs do not represent the underlying function correctly, errors in use occur when the user's model does not comply with the computer's model. Thimbleby (2018) puts it simply: "…both user and computer "do things" but they do not necessarily do the "right things" from each other's point of view. Then a sequence of misunderstandings ensues, which will end in frustration for the user" (p. 13). Those model mismatches do not cause disasters most of the time, rather, they usually occur as irritations that users can get away with ignoring, working around, or dismissing (Thimbleby, 2018). However, the costs of these errors, when encountered in fields such as health or the military, are much greater than those encountered in products used in everyday life.

When we consider the fact that users are not homogenous, they have a different range of skills, knowledge, and preferences, we can assume that fixed interfaces will be better suited to some users than to others (Benyon & Murray, 1993). The primary objective of user modeling is to adapt the interfaces to users to make the use of a system more inclusive by facilitating interaction. For example, a user modeling system called AVANTI (Stephanidis et al, 1998) was proposed to provide an accessible and high-quality user interface to all potential end users. The system employed adaptability and adaptivity techniques that enable to tailor the user interfaces according to the user's abilities, skills, requirements, and preferences (Stephanidis et al, 1998). Adaptability refers to the process of modifying aspects of the interface according to static user characteristics that are known before the interaction and that are assumed to remain unchanged during the session such as "*physical abilities*" ("whether the user is able-bodied, blind, or motor-impaired"), "*the language",* "familiarity of the user with computing, networking, hypermedia applications", "*the overall interaction target*" ("speed, ease, accuracy, error tolerance"), and "*user preferences*" (Stephanidis et al, 1998, p. 159). On the other hand, adaptivity refers to the process of modifying aspects of the interface according to dynamic user characteristics and situations detected at run-time such as "*user familiarity with specific tasks*" ("capability to successfully initiate or complete tasks"), "*ability to navigate*", "*error rate"*, "*disorientation*", "*user idle time"*, and "*repetition of interaction patterns*" (Stephanidis et al, 1998, p. 160).

In addition to adapting the interface according to the end user, user models also make it possible to evaluate different design alternatives before implementation by simulating users' interaction (Biswas & Springett, 2018). Predictive cognitive models like GOMS or KLM are widely used to estimate the duration of task completion (Biswas & Springett, 2018). The existence of user models suitable for testing the product may provide an advantage to developers by easing the dependency on user participation in design research. However, designers should not just analyze and predict how the user will interact with the interfaces, they should also validate these predictions by testing the behavior of real users in the real world (Liu et al, 2010).

Although it is thought that existing models are incapable of encompassing to elucidate every kind of interaction, one thing is for sure they have served as a foundation for developments to better understand the relevant issues and design priorities (Biswas & Springett, 2018). Thanks to these models, we conceived that more aspects of humans are relevant to interaction; that users cannot be defined as one type, and that the division of users cannot be reduced to only two definitions: novice and expert. We now pay more attention to being inclusive.

2.1.3.4 User profiling

As a widely used term in the industry, the user profile is used as a method to represent real users based on existing data. It is almost the same as a persona like a fictitious person that represents a collection of attributes (Kuniavsky, 2003). The purpose of profiling users is to help designers better understand the actual user, especially to be used in the early stages of product development and as a basis for usability evaluation (Liu et al, 2010). A good understanding of users requires user analysis which often means distinguishing their characteristics such as age, gender, expertise with technology, educational background, attitude toward technology, linguistic ability, etc. (Liu et al, 2010). The impact of such characteristics on the use was stressed by many researchers (Botwinick, 1973; Welford, 1977; Owsley et al, 1991; Broos, 2005; Faulkner & Wick, 2000; Preece, 2002; Janhager, 2005).

The age of users is an important determinant in interaction that needs to be considered by designers. As people get older, the physiology of the eye changes (Darin et al, 2000); their visual attention decreases (Owsley et al, 1991); it takes longer time to complete computer-based tasks and they make more errors (Czaja & Sharit, 1993); and they experience fear of unintended consequences of using unfamiliar technology (Kurniawan, 2008). Older users experience significantly higher anxiety than young users while using computers (Laguna & Babcock, 1997).

Some studies have suggested that gender can be a determinant influencing interaction. Philbin et al. (1995) investigated the differences in learning styles between men and women by including 72 participants from different ethnic backgrounds. They found that while the issue of "*concern for others*" was the more prominent response among women, "*concern for self*" was the primary male response and that while women learned better in hands-on and practical settings, men were good at learning abstract concepts (Philbin et al, 1995, p. 490). Barret and Lally (1999) investigated the gender difference in learning and socio-emotional behavior in a formal online learning environment by analyzing the content in online dialogues and they found that males sent more and longer messages than females. Broos (2005) aimed to reveal the gender divide in information and communication technologies by including 1058 participants and reported that males were found to have less anxiety than females, that respondents who have used computers for a longer period showed less anxiety, however, while the experience had a positive impact on decreasing anxiety for men, a similar effect was not found for women. It was observed that boys had significantly more positive affective attitudes toward computers than girls (Levin & Gordon, 1989; Martin, 1991; Corston & Colman, 1996; Sherman et al, 2000). On the other hand, some studies found no gender difference in attitudes toward simple computer tasks (Busch, 1995), in interest in computer use among respondents with the same level of experience (Chen, 1986), and in computer skills (Hatlevik & Christophersen, 2013; Van Deursen & Van Diepen, 2013; Ihme & Senkbeil, 2017).

The cross-cultural design of interfaces is currently pervasive, which brings up concerns about its validity and the impact of cultural differences on users' behavior in interaction (Liu et al, 2010). An experimental study (Noiwan & Norcio, 2006) investigated the effects of animated colors on the attention and perceived usability of users from two cultural groups, American and Thai. They suggested the influence of cultures on overall performance, retention, and self-reports on usability (Noiwan & Norcio, 2006). Dong and Salvendy (1999) investigated the implications of different languages on the menu designs in interfaces and found that while horizontal English menus resulted in better performance than vertical ones, vertical Chinese menus resulted in better performance than horizontal ones. Therefore, they concluded that interface layouts should be congruent with the language being used. On the other hand, Ford and Gelderblom (2003) investigated the effects of Hofstede's cultural dimensions (Hofstede, 1991) on human performance achieved by a computer-based user interface and found no evidence that such an effect exists. A study aimed to compare the computer anxiety of Turkish and Dutch university students and found that Turkish students experience significantly higher anxiety than Dutch ones independent of gender (Tekinarslan, 2008). However, they stated that the differences in computer anxiety between the two groups could be explained by experiences rather than cultural factors since Dutch students had higher computer ownership (Tekinarslan, 2008).

Studies of HCI focusing on the cultural influence in interaction are quite immature. Culture remains difficult to study alone, especially in relation to HCI since people's behaviors might be influenced by other factors rather than cultural traits (Smith & Yetim, 2004). There is limited evidence to the contrary and current research is unable to establish definitive methodologies for the design of usable systems in specific cultures (Smith & Yetim, 2004). Clemmensen and Roese (2010) analyzed studies on the culture that were published in journals between 2000 and 2010 and found that most studies adopted quantitative methods (20/27), possibly due to the need for local experts and increased costs for qualitative methods; that most studies included students (14/27), that more than half of these studies used questionnaires and only four were fieldwork. In that sense, the fact that the majority of studies investigating culture use quantitative methods supports this argument.

Some studies have evaluated the effect of culture on methods used in usability studies in this field. The impact of cultural differences on usability evaluations was discussed by adopting the view of cultural cognition (Nisbett et al, 2001) that focuses in particular on two broad groups of people – “Easterners” and “Westerners” (Clemmensen et al, 2009) and showed that think aloud method may be affected by culture. Similar results were found in a study that examined the influence of the interviewers’ cultures on the results of interviews conducted for usability (Vatrapu & Pérez-Quiñones, 2006). The study showed that participants found more usability problems and made more suggestions to an interviewer who was a member of the same culture than to a foreign interviewer.

A relatively recent study that examined the usability of interfaces for middle-aged learners showed that educational background was the major factor influencing the computer use of the participants (Chou & Hsiao, 2007). They found that the low-educated participants struggled with using the mouse, reading information on the monitor screen, memorizing, and comprehending the content, using a software interface, and solving error problems (Chou & Hsiao, 2007). The effect of computer experience on interaction is self-evident, which may be closely related to educational background as well. Tekinarslan (2008) reported that participants who own computers are less likely to experience anxiety during the interaction.

2.1.3.5 A model proposition for profiling the users of expert systems

In the section on defining the user, I have classified the user through three main attributes in the context of expert systems, which were the user's level of knowledge (Whether the user had undergone training in the relevant field was considered a variable in relation to the knowledge), the user's experience, and the user's competence. In this section, I will examine these parameters in more detail and propose a 3D model for user profiling in the field of expert systems.

As mentioned in the previous section, experience is an important parameter in interacting with systems, and gaining experience is closely related to being exposed to systems. In the expert system context, while I define "novice users" as those who have never used expert systems before and do not have any experience, I define "advanced users" as those who are experienced in expert systems or accustomed to similar software. It has been mentioned in the literature how the participant experience should be considered in terms of methodology in research design and how users from different experiences can affect the results (Chapanis, 1959; Johnson & Baker, 1974; Engelbrektsson et al, 2000). Fisher and Kaplan (1990) propose a checklist of items that would enable qualitative and quantitative comparisons by using descriptive statements for selected users in research design. These are as follows:

- Duration of use (overall period of time using similar tools);
- Frequency of use (e.g. number of times per week);
- Extent of user sessions (time per session);
- Flexibility of task goals (fixed/routine or flexible/creative involvement);
- The number and type of application package or operating system used (breadth of experience);
- The relative length and type of formal training or instruction in computer science, programming, etc.

Taking some of the items that Fisher and Kaplan (1990) suggested one step further, I offer some axes and variables which would help researchers define the profile of the user who uses an expert system.

As briefly mentioned in the first section, there are 3 main axes: user's knowledge, experience, and competence which are illustrated as a graph in Figure 2.7. On the graph, "*x*", "*y*" and "*z*" respectively represent the user's experience, competence, and knowledge. The minimum value of these axes is zero because there is no such thing as minus experience, competence, and knowledge. There is a limit on the z-axis for the non-expert user. Because no matter how interested the user is in the field (in the medical context, sometimes the user can be suffering from a chronic illness or a relative of the patient), the decisive factor here is whether the person has undergone formal training (T), which makes the user an expert. Therefore, I arbitrarily define the limit (L) as an indicator of the maximum level of knowledge a non-expert user can obtain.

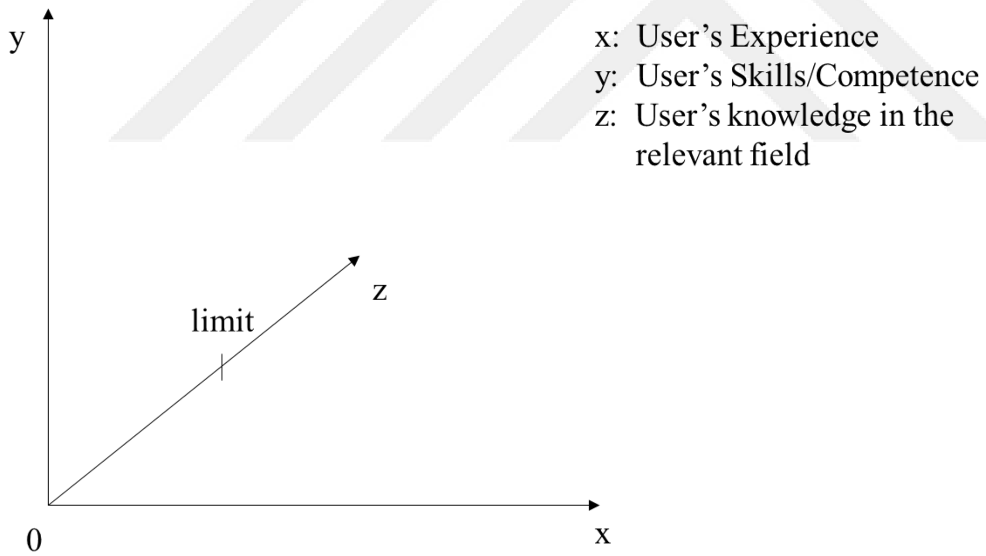


Figure 2.7: Three axes of determinants in defining a user.

Other variables can be listed as follows:

d: Duration of use. (How long has the user been using the expert system?)

f: Frequency of use. (How often does the user use the expert system?)

s: Extent of sessions. (How much time does the user allocate for each use?)

w: User’s willingness. (How extensively has the user the willingness to operate the system?)

The user's experience is dependent on how long (d) and how often (f) they have used the system, and how much time (s) they allocated for each session. The user’s willingness (w) and the user’s competence (y) also influence the extensity of the experience. For example, if a user is enthusiastic about learning the system but has a visual impairment, her/his experience (x, d, f, s) will be affected by these parameters (Figure 2.8).

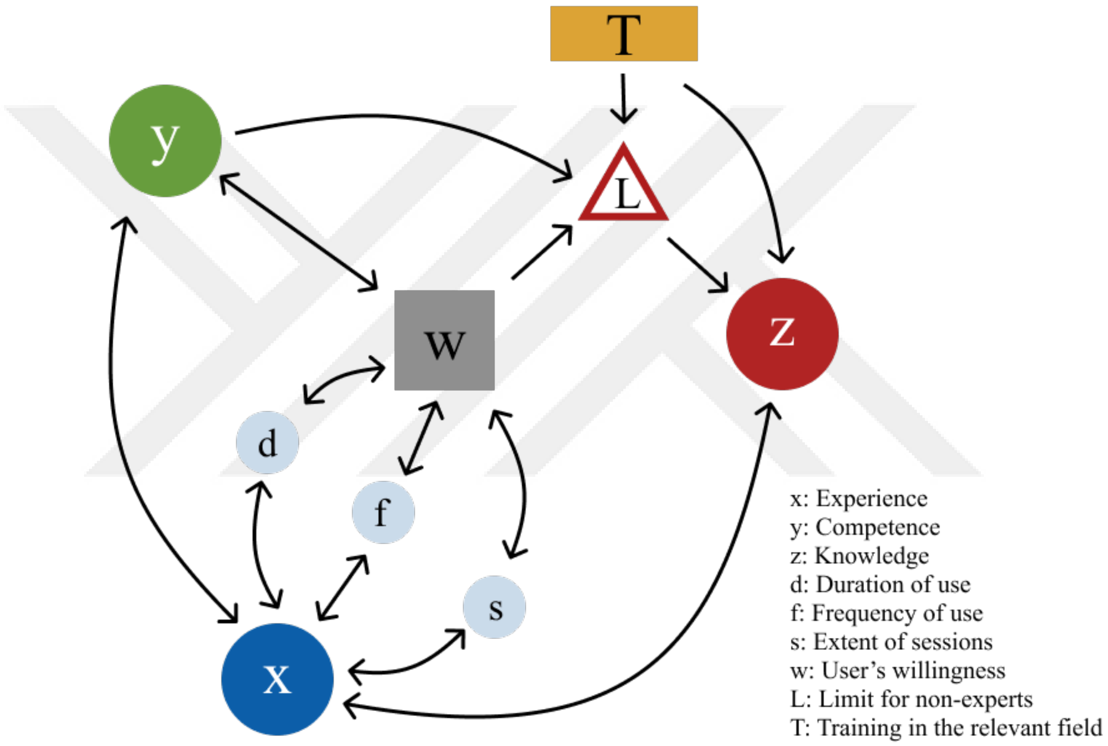


Figure 2.8: The relation between variables.

The user’s competence (y) is also influenced by the experience (x) because the more often the user uses software and uses it for a long time, the more they may become adept (Figure 2.8). However, there is also a limit that is determined by the “*internal factors*” (Osvalder & Ulfvengren, 2009).

The user’s willingness (w) may be influenced by the experience (x). For example, previous frustrations experienced by the user with such systems may reduce the user's desire to use other systems, which may negatively affect the frequency of use (f), and the extent of sessions (s), hence the possibility of future experience. The user's competence (y) may also influence the willingness (w) to use a product (Figure 2.8). If the system is found too complex by users, they may experience anxiety, which will reduce their motivation to use it. We often see this case in elder users (Laguna & Babcock, 1997).

As mentioned above, there is a limit (L) to the knowledge a non-expert user can obtain regardless of their willingness. Being an expert requires training, which is the main condition for the expert's deep knowledge in the field. While the limit for how deep non-experts can acquire knowledge is roughly clear, where it sits within a range is determined by the user's competence, experience, and willingness to learn. The limit does not apply to expert users. The level of knowledge (z) affects the experience (x). For example, a novice expert user may use an expert system easier than a non-expert user due to being familiar with the content. Reciprocally, a non-expert user's knowledge can increase by gaining experience with the system.

So, what makes this model different than others? First of all, it is a novel proposal for the literature since there is no model in which the user is defined, and profile properties are determined according to expert systems. Unlike others, I added the user's competence to the model and parsed the competency and willingness to use the system extensively parameters differently than Fisher (1991). Defining users as non-experts and experts according to the system is new because the term expert has different uses in the literature, where it sometimes refers to experienced users (Allwood, 1986), and sometimes it refers to users who are motivated to use the system extensively (Fisher, 1991). As a new concept, I defined an arbitrary limit in knowledge for non-expert users.

This model added a third dimension to two-dimensional classifications and introduced new parameters affecting the axes (Figure 2.9). The model can be used to categorize participants in research on users of expert systems and can be useful for designers to create personas during the design phase of the system. It should be noted that this is only a proposal and is not yet supported by empirical data.

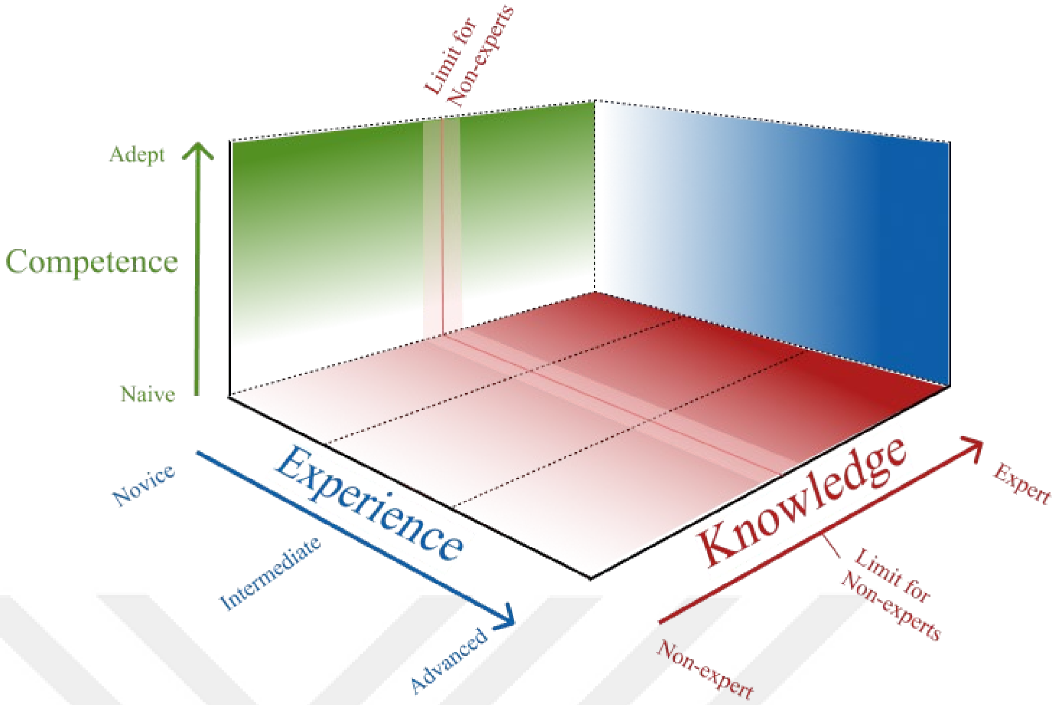


Figure 2.9: 3D model of user profiling using an expert system.

2.2 Digital Oral Health Applications

Digital health products hold great potential for improving public health by providing a wide range of functions such as enabling self-diagnosing, promoting self-monitorization of chronic diseases, increasing users' health literacy, acting as reminders for medications, facilitating communication between patients and healthcare professionals, enabling remote treatment modalities, making behavioral interventions like smoking cessation, etc. According to a report published in 2022 (Turner, 2022), more than 6,5 billion people own smartphones which translates to 83.32% of the world's population. Thanks to their widespread use, mobile apps play a major role in making these functions accessible for people from different educational and socioeconomic backgrounds. Numerous publications show the impacts of digital technologies on clinical outcomes, treatment courses, symptom management, and the patient's quality of life scores in different subfields within medicine such as psychiatry (Graham et al, 2020; Barnett, 2018), oncology (Larson et al, 2018), cardiology (Kotecha et al, 2018), internal medicine (Whitehead et al, 2016), infectiology (Daher et al, 2017) and gynecology (Scott et al, 2015). However, while some studies (Agarwal et al, 2019; Scheibe et al, 2015) have been unable to demonstrate significant clinical benefits due to low user engagement to applications, some (Yin et al, 2019) found mixed results. There are doubtless more aspects to be scrutinized such as usability, user acceptance, and design before drawing any conclusion. Despite the lack of standardization in evaluations and inadequate clinical data, mHealth technology is still seen as a promising tool to improve healthcare and patient empowerment.

Mobile app technology also offers great opportunities in the oral health field. This is a relatively new concept called mDentistry (Xiao et al, 2021) aims to improve public oral health and oral health services. The mDentistry apps provide many advantages to both patients and dentists with functions like enabling patient-provider interaction, remote screening, treatment planning, and self-diagnosing, improving oral hygiene practices, educating patients and caregivers, and guiding in dental emergencies. However, when compared to mHealth tools, their use and number are quite limited. Accordingly, there are also a limited number of studies assessing the body of evidence on these applications.

2.2.1 Approaches in the literature

The studies on oral health applications in the literature mostly aim to investigate their benefits and effectiveness. According to a recent systematic review (Qari, 2019, p. 25), over half of the included 38 studies were published in dental journals, while nearly a quarter were published in health informatics and mostly in Telemedicine journals. As mentioned before, there is no design study evaluating those apps. Nonetheless, some studies in the literature adopt design methodologies to review the apps or use the user-centered design approach in the development processes of those apps.

2.2.1.1 Clinical approach in mDentistry

Oral health is an essential part of our overall health and well-being. It plays an important role in our daily functions such as eating, speaking, smiling, and socializing. Any problem in that area causing pain and discomfort may hinder our routines and impact our quality of life. Poor oral health is not only associated with physical health but also negatively affects self-confidence and emotional state, leading to undesirable social life consequences. On the other hand, good oral health reflects an individual’s ability to maintain self-care and show importance to their health. People with an ideal smile may have a better chance of getting a job than those who don't (Pithon et al, 2014).

Most oral diseases have a multidimensional etiology including environmental factors and genetic predispositions. Poor oral hygiene, high sugar intake, low fluoride exposure, low health literacy, tobacco use, and malnutrition are the most common driving factors (WHO, 2022). Like many diseases, oral diseases are closely related to the lifestyle and daily habits of the individual. Therefore, experts often recommend brushing with fluoridated toothpaste, regular dental visits, and avoiding frequent sugar consumption and smoking.

Although oral diseases are largely preventable, they are highly prevalent. According to Global Burden of Disease reports (Bernabe et al, 2020), they affected nearly 3.5 billion people in 2017, and dental caries is the most common disease. In 2010, severe periodontitis (the oral disease affecting tissues supporting teeth) was the sixth-most prevalent health condition, affecting 10.8% of people worldwide (Kassebaum et al, 2014). Oral cancers were among the top 15 most common cancers in the world in 2018 (Bray et al, 2018). Dental caries and periodontal diseases cause tooth loss which is strongly correlated with low quality of life (Gerritsen et al, 2010). Poor oral health may even hinder school performance in children (Anandakrishna & Chandra, 2012).

Oral diseases affect systemic health as well. Numerous studies report the negative effects of periodontal diseases on cardiovascular conditions (Mattila et al, 1989; Umino & Nagao, 1993; Beck et al, 2001, 2005; Hung et al, 2004; Engebretson et al, 2005; Desvarieux et al, 2005; Pussinen et al, 2004), adverse pregnancy outcomes like preterm and low-weight births (Offenbacher et al, 1996; Davenport et al, 2002; Goepfert et al., 2004; Radnai et al., 2004; Jarjoura et al., 2005), diabetes (Cianciola et al, 1982; Thorstensson & Hugoson, 1993; Safkan-Seppala & Ainamo, 1992; Shlossman et al, 1990) and respiratory infections (Hayes et al, 1998; Scannapieco et al, 2003; Azarpazhooh & Leake, 2006). Two meta-analyses suggest a positive correlation between periodontal disease, edentulism (toothlessness), and the risk of oral, lung, and pancreatic cancers (Michaud et al, 2017; Maisonneuve et al, 2017).

Oral diseases place significant economic burdens on societies (Listl et al, 2015). Besides their intangible costs like pain and discomfort with biting, chewing, eating, and smiling, there are direct (treatment expenditures) and indirect (productivity losses due to absence from work or school) costs (Peres et al, 2019). The findings of a study indicate that the 2015 global economic burden of dental diseases is a total of $544.41 billion including direct and indirect costs and the expenditures related to general health problems that occur as a result (Righolt et al, 2018). According to Health Statistics Yearbook (2018), total oral health expenditure in Türkiye is approximately $1.65 billion. While the highest dental expenditure per capita is in North America, Australasia, and West Europe, the lowest is in Sub-Saharan countries (Righolt et al, 2018).

The term "*social determinants of health*" was defined by the World Health Organization (2016, as cited in the American Academy of Pediatric Dentistry (AAPD), 2020b) as "the conditions in which people are born, grow, work, live, and age and the wider set of forces and systems shaping the conditions of daily life" (AAPD, 2020b, p. 28). Like general health, social determinants of oral health are also greatly influenced by socioeconomic status. Social inequality and poverty pose a considerable risk for oral diseases as they are associated with limited access to dental healthcare and oral hygiene products, inability to purchase healthy food, low educational status, etc. This causal relationship was highlighted by a study that included tsunami survivors to investigate the impact of economic deterioration on the probability of tooth loss (Matsuyama, 2017). Some vulnerable communities like indigenous populations in high-income countries (Jamieson et al, 2016), refugees (AAPD, 2020a), and children from low-income families (Antunes et al, 2006) are at high risk for dental caries and tooth loss.

Lately, attention has focused on mobile application technology to improve public oral health, promote preventive measures, and alleviate social inequality. Making oral disease screening services, self-diagnosing tools, and scientific oral health information accessible via mobile apps to all citizens regardless of their socioeconomic background is one way to reduce the disparity. However, there are a limited number of applications on the market to meet this need and the standards in terms of scientific content and usability. Therefore, WHO recently (2021) published an implementation guide for app developers and innovators. The guideline introduces four key modules that address a range of target groups. These modules are: "*mOralHealth literacy*" to improve the oral health literacy of individuals and communities and raise awareness, "*mOralHealth training*" to enhance the knowledge and skills of general and oral health professionals and encourage them to use mobile technologies for e-learning, "*mOralHealth early detection*" to provide information on the potential of early detection by using remote diagnostics tools and to improve timely treatment and quality of health care services, and "*mOralHealth surveillance*" for collection of epidemiological data as well as monitoring of quality patient care and service delivery.

Oral health education is an essential tool for raising people's awareness of the etiology, risks, and precautions of oral diseases. As mentioned above, one of the most common causes is poor oral hygiene. Tooth brushing at home is the most widespread plaque control method and motivation plays a key role in this practice. Studies showed that verbal instructions by dentists are not sufficient to make a significant change in patients' toothbrushing habits, but written and visual instructions are more influential (Ay et al, 2007). The fact that mobile applications make visual content easily accessible has created a potential alternative in oral health education. Most studies in dentistry literature focus on evaluating the effect of educational interventions through mobile app technology. While some studies attempted to show the clinical significance of these interventions by using already existing messaging apps (WhatsApp, WeChat, etc.) or the apps specifically designed for that purpose, others only assess their impact on the level of oral health knowledge. The effectiveness of sending educational and reminder messages to adolescents under orthodontic treatment in increasing compliance and attendance to regular visits, reducing bracket bond failure and thereof treatment duration (Li et al, 2015), and improving oral hygiene parameters (Zotti et al, 2015; Bowen et al, 2014) was shown. A study comparing the benefit of sending educational content via WhatsApp and PowerPoint presentations on tobacco use and oral cancers showed results in favor of messaging (Nayak et al, 2017). The apps developed for different ranges of age groups and patients were screened and they were found effective in improving overall oral hygiene status and oral health awareness in the elderly population (Lee et al, 2021) and adolescents under orthodontic treatment (Alkadhi et al, 2017; Scheerman et al, 2018), improving knowledge of oral hygiene practices in

10–12-year-old female students (Rahaei et al, 2021), knowledge of pediatric oral health in mothers (AlKlayb et al, 2017; Alqarni et al, 2018) and understanding orthognathic surgery and its associated complications (Pulijala et al, 2016). A systematic review (Toniazzo et al, 2019) analyzed the results of published studies and reported that the use of educational apps and sending text messages related to oral hygiene and/or oral health improves knowledge, raises awareness, and reduces clinical parameters such as plaque indices and gingival inflammation.

Increased oral health awareness does not necessarily mean good oral hygiene. Technique and duration of tooth brushing are critical in maintaining good oral hygiene, so keeping brushers motivated until the efficiency reaches the optimum level is necessary. Brushing is more effective in plaque removal when its duration is longer than 2,5 minutes in preschoolers (Pujar & Subbareddy, 2013). It can be difficult to motivate a child to brush for that long. Maintaining optimal oral hygiene in orthodontic patients, the majority of whom are adolescents, is even more challenging, and brackets complicate this further. As highlighted before, interactive tools with visually appealing materials are more convenient to sustain motivation. Many oral-health-oriented mobile apps aim to assist individuals while brushing their teeth to ensure adequate time is spent on each side of the dentition by using animations and other visual features. There are different types of apps assisting toothbrushing on the market that can connect to toothbrushes, not only via Bluetooth but also external sensors, and that act as a timer without being connected. Oral-B® developed a brushing app that can connect to its electric toothbrushes, calculate brushing duration, detect the pressure exerted by the brusher, and verify the cleanness of all the mouth areas by using the smartphone camera. A randomized controlled study investigated the efficacy of using this app and the electric toothbrushes together in adolescents with fixed orthodontic appliances (Deleuse et al, 2020). It was reported that there was no significant difference between the test and control groups in terms of oral hygiene and that there was a decrease in the use of the application over time (Deleuse et al, 2020). Another study investigated the effect of the manual toothbrush attached to a sensor that connects to the app, tracks the brushing movements, and adds bonus points each time the child gets to optimum brushing movements (Alkilzy et al, 2019). It has been shown that there was an improvement in the oral hygiene status of preschool children who use the toothbrush with the sensor and the app (Alkilzy et al, 2019). The only oral health app approved by the National Health Service of the United Kingdom "Brush DJ" was utilized in a study to test its efficacy on clinical parameters in orthodontic patients (Farhadifard et al, 2020). This app aims to prompt users to brush for 2min while listening to music played from the user's device or streaming services, reminding them twice a day and including links to educational YouTube videos that show brushing techniques and how to use other oral hygiene appliances (Farhadifard et al, 2020). Brushing frequency and duration were positively correlated with the use of the app, so the patients in the test group showed better oral hygiene performance (Farhadifard et al, 2020).

Oral health behavior is established in childhood. For this reason, children need to start brushing as soon as their teeth erupt and learn which behaviors are harmful to oral health. Some apps aimed at improving children's oral self-care and behavior change adopt the gamification approach. The most widely accepted definition of gamification is "the use of video game elements in non-gaming systems" (Edwards et al, 2016, as cited in Zolfaghari et al, 2021). Although there is no consensus, an increasing number of studies suggest that gamification is effective in the healthcare field (Cugelman, 2013). The effect of gamified oral hygiene apps on improving toothbrushing quality in children has been shown in several studies (Jacobson et al, 2019; Zolfaghari et al, 2021).

With the expansion of internet connection coverage and the improvement of the camera features of smartphones, health services have become remotely available through mobile applications. The oral health service in which communication is provided via mobile phone or computers, where the dentist is at one end and the patient at the other end is defined as teledentistry. It is the use of health information technology and telecommunications for oral care, consultation, education, and public awareness with the broad goal of improving oral health (Daniel & Kumar, 2014). Health experts benefited from teledentistry for cancer screenings in regions with challenging conditions. The accuracy of diagnoses through apps that were developed for this purpose was investigated (Gomes et al. 2017; Birur et al, 2015; Uthoff et al, 2018). With the help of apps, frontline health workers were able to be coordinated by the specialists in the centers, and remote oral cancer surveillance in rural places where the infrastructure is insufficient became possible. Another study showed that teledentistry had the potential to detect caries (Estai et al, 2016, 2017). The researchers found promising results but underlined that the diagnosis made through the applications has not yet reached the gold standard and there is a need for advances in technology.

Management of trauma-related dental emergencies at the accident site has critical importance in prognosis and the evidence was presented by many studies (Chappuis & von Arx, 2005; Andreasen et al, 2006; Schwartz et al, 2002; Andersson et al, 2012; McIntyre et al, 2007). Physically active people such as school children or athletes are at high risk for dental trauma. Teachers, sports instructors, or people who do sports as amateurs or professionals should be aware of the steps they have to follow for proper care in the event of an accident. A study that assessed the knowledge of dental emergency management in a group of schoolteachers and compared different educational interventions showed that the app (Dental Trauma®) is an effective means of providing accessible information (Al-Musawi et al, 2016). Another study attempted to determine the efficacy of a clinical decision-support tool for dental trauma and compare the effects of its print and mobile app format by including novice and expert clinicians and showed that the app was more effective in improving diagnosis than the print format (Machado et al, 2018).

Some studies assess dentist-centered apps that were developed for different purposes. A study aims to compare the cephalometric evaluation (the special radiographic assessment done before orthodontic treatments) results of conventional tracing with those acquired through app tracing and found significant differences between the two methods, which is not desired but noted that tracing with the app was faster (Sayar & Kilinc, 2017). Some focus on the efficiency of the apps developed for dentistry students to facilitate learning dental anatomy (Javaid et al, 2016; Rodrigez-Andres et al, 2017), prosthodontics (Deshpande et al, 2017), and dental anesthesia (Ribeiro et al, 2017).

As mobile applications take more place in our lives, we see more about oral health, and therefore researchers evaluate these applications and make more publications. However, there are limitations regarding these studies that scrutinize their clinical benefits. They show high variabilities such as follow-up time, risk of bias, and different indices to assess dental plaque and marginal gingival bleeding, lack of standards in methodology, and a narrow range of participant age, which may influence the results (Toniazzo et al., 2019; Fernández et al., 2021). Other factors hinder conducting comprehensive studies such as the difficulty of maintaining user engagement with the app for a longer period. Therefore, evaluating the effectiveness of these technologies, especially their persuasive aspects is a great challenge (Kientz et al, 2010).

2.2.1.2 The categorizations of oral health apps in the literature

While some reviews evaluated the oral health applications on the market, some examined the published studies focused on those apps to categorize them according to their functions, themes, intervention methods, and target users.

Tiffany et al. (2018) screened 1975 Android and 1005 IOS apps related to oral health and filtered them according to some selection criteria such as excluding non-English, and those designed for dental professionals, dental students, patients of specific dental practices, or children. They included the 33 most popular oral health apps based on the assumption that these were likely the highest quality (Tiffany et al, 2018). It was found that two-thirds (22/33, 67%) targeted the public, while some (4/33, 12%) were offered to be used as an adjunct for their electric toothbrushes (Tiffany et al., 2018). The key functions of those apps were listed as follows: timers for tooth brushing (20/33, 61%), oral hygiene alerts, and reminders (13/33, 39%), oral health educational content (13/33, 39%), tracking toothbrushing (9/33, 27%) and dental appointments (7/33, 21%), ability to search for dentists (5/33, 15%) and communicate with them (5/33, 15%) (Tiffany et al, 2018).

Another review screened 562 apps and only included those (N=40) in English that are associated with dental caries prevention factors addressing oral hygiene practices, fluoride usage, and diet (Chen et al, 2021). It was shown that the apps were geared toward adults or young adults (22/40, 55%), adolescents (14/22, 63%), and children (11/40, 28%) and that almost 93% of them (37/40) addressed oral hygiene, 43% of them (17/40) fluoride use, 53% of them (21/40) diet (Chen et al, 2021).

A systematic review analyzed 37 oral health apps designed for patients and healthcare providers in 38 peer-reviewed studies with empirical results and presented that the oral health topics discussed in those studies were respectively about oral pathology including caries, periodontal diseases, dental injuries, cleft lip, and oral cancer (37,8%), oral hygiene behavior (29,7%), dental anatomy (24,3%), oral health knowledge (18,9%) to educate patients and caregivers on infant and child oral health, dental treatment procedures (16,2%), and dental traumas (10,8%) (Qari, 2019, p. 28). Some apps consist of more than one theme, so the sum of the percentages is greater than 100% (Qari, 2019, p. 28). These apps target health professionals including dentists and doctors (54%), allied health providers such as dental hygienists and therapists (35,1%), dental students (10,8%), and the public (35,1%) (Qari, 2019, p. 31). The purposes of these apps are shown in Venn diagrams (Figure 2.10). Qari (2019, p. 131) conducted another analysis to assess the no-cost, stand-alone oral health apps (N=52) in English only targeted the patients and found that the majority (16/52, 31%) is related to oral hygiene behavior such as toothbrushing timers and the ones connected to electric toothbrushes, 15% (8/52) aims to improve oral health knowledge, 11% (6/52) enables the communication between patients and dentists, %9 (5/52) provides home remedies for oral diseases, the rest is about dental traumas (3/52, 6%), orthodontics (3/52, 6%), bad breath (1/52, 2%), oral mucositis (1/52, 2%) and oral cancer (1/52, 2%).

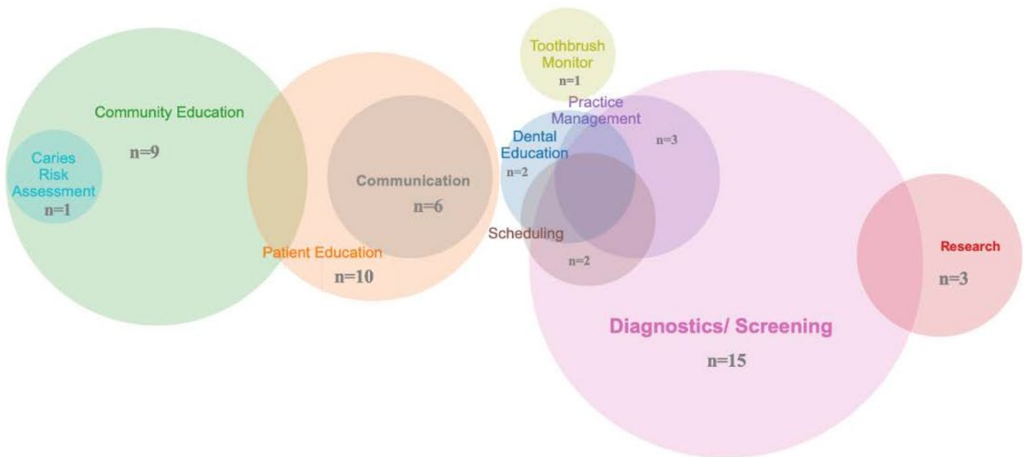


Figure 2.10: Purpose of mobile dental apps published in the peer review literature (Qari, 2019, p. 33).

According to an article that examines the mHealth industry and medical device regulations, which will be scrutinized in the following section, medical apps can broadly be categorized under four types (Roth, 2014): (1) those providing medical information to users (laypeople or healthcare providers), (2) those functioning as diagnostic tools, analyzing inputs and computing results, (3) those enabling users to control unattached medical devices, and (4) those that function in cooperation with sensors, or other devices attached to smartphones and act as adapters (Roth, 2014).

Although there is no clear distinction between the third and fourth categories described in the article in terms of their dependency and degree of control authority, the attached apparatuses appear interdependent with the apps that act as adapters and they do not work without each other, unlike the unattached medical devices. One can ask whether the apps connected to electric toothbrushes fall under the third category. Because these apps are unable to freeze the function of electric toothbrushes.

2.2.1.3 Design approach in the evaluations by health experts

The research framework of oral health applications is expanding due to the need to be treated as a product. When it is used by general users, there are always other parameters worth to be investigated. Therefore, in addition to studies that investigated the clinical effects of oral health apps available on the market or specifically developed for research, some reviews in the literature attempted to analyze their functional features, usability, the theoretical basis of the content, user satisfaction, and behavior intervention models by using various methodologies. These reviews were carried out by either the authors or independent experts included in the studies.

The widely utilized tool in the reviews is the Mobile App Rating Scale (MARS) developed as a reliable, multidimensional measure for trialing, classifying, and rating the quality of mobile health apps (Stoyanov et al, 2015). The scale consists of 6 main sections: the first section aims to reveal the features in terms of engagement, interaction, and entertainment; the second section focuses on their functionality, ease of use, and navigation; the third section includes questions about aesthetic appeal, visual and stylistic consistency; the fourth section helps to scrutinize the quality of the information provided, and the consistency of the app description in the app store with real life; the fifth section and the final section have scorings about subjective statements and evaluator’s opinion on the possible impact of the app. Each item in the sections is scored using a five-point Likert scale (1: inadequate, 2: poor, 3: acceptable, 4: good, and 5: excellent) and the total scores are averaged at the end of the evaluation.

A review that evaluated 40 dental caries prevention apps available in Australian App Stores found that less than half of the apps were of good quality according to their MARS score, and oral hygiene apps scored high because they had a greater number of features (Chen et al, 2021). Two studies using a similar method assessed 15 oral health apps for children available in sub-Saharan Africa (Kanoute et al, 2022) and 10 oral hygiene apps available in France (Carrouel et al, 2022) and displayed the mean scores on a heat map. Another study aimed to assess 20 oral hygiene apps using the MARS tool reported that the apps appeared to be functional, but an improvement in engagement, aesthetics, and information content is needed (Sharif & Alkadhimi, 2019).

Another method used to assess oral health apps is the "Persuasive Health Technology Heuristics" (Kientz et al., 2010) which was a modified version of Nielsen's (1994) "10 Usability Heuristics" rule of thumb in interaction design. Kientz (2010) describes 10 heuristics for persuasive health technologies as (1) appropriate functionality, (2) not irritating or embarrassing, (3) protecting users' privacy, (4) use of positive motivation strategies, (5) usable and aesthetically appealing design, (6) accuracy of the information, (7) appropriate time and place, (8) visibility of user's status, (9) customizability, and (10) educating users. Since some oral health apps aim to change users' behavior in terms of oral hygiene practices, diet, etc., Tiffany et al. (2018) conducted an analysis using "Persuasive Health Technology Heuristics" with a scoring system to assess basic descriptive characteristics, evidence of theoretical basis, functionality, usability, and message framing of 33 oral health apps. They found that none of the apps were empirically validated; more than half of them included written content intended to change behavior as opposed to basic instructions to show users how to use features; the majority utilized both gain- and loss-framed messages to influence the user behavior; and they are scored poorly on protecting privacy, and the overall quality (Tiffany et al, 2018).

A recent review (Ho et al, 2022) analyzed 54 apps for promoting children's oral health by using the "Health on the Net Foundation Code of Conduct" (HONcode) and the "Scientific Basis of Oral Self-care" (SBOSC) guideline. HONcode consists of 8 criteria: (1) authority (disclosed information about the developer team), (2) complementarity (underlining the limitation of the app which does not replace the doctor-patient relationship), (3) confidentiality (privacy of personal data), (4) validity (the content having “last updated” date), (5) justifiability and objectivity (health information with references), (6) user’s practice (app’s user-friendliness and reachable team members), (7) financial disclosure (a transparent funding sources), and (8) advertisement policy (ads that are clearly separated from the health content) (Health on the Net, 2020). On the other hand, the researchers derived 6 factors from the SBOSC guideline (NHS Health Scotland, 2011, as cited in Ho et al, 2022), which are (1) choice of toothbrush (size and bristle), (2) use of fluoride toothpaste (fluoride concentration and amount), (3) brushing time (frequency, timing, and duration), (4) adult supervision, (5) brushing techniques, and (6) other advice for the prevention of caries. They observed that the new applications are of better quality in terms of scientific basis than the old ones, but this quality is not sufficient in general (Ho et al, 2022).

2.2.1.4 The design methods adopted in the development processes of oral health apps

Besides the systematic reviews that analyze the oral health apps available on the market, there are also studies in the literature presenting the design methods adopted at different stages of the development process. These studies were conducted with different motivations such as collecting data for design refinement for the apps in the prototyping phase or for demonstrating user satisfaction of those in the final stage. A variety of methods were used in these studies to investigate the usability including target users at different stages by collecting data at different depths.

A recent case study aimed at refining and testing the usability of an AI-powered oral health app (AICaries) prototype utilized the data gathered from target users (parents with young children) for the app’s design (Xiao et al, 2021). The app’s AI enables parents to detect caries in their child’s mouths through its image recognition technology (Xiao et al, 2021). However, the study did not aim to assess the AI’s performance, but the feedback about its usability, difficulties experienced by parents, especially when taking pictures of their children's mouths, and user flow (Xiao et al, 2021). There were two steps in the study: in the first one, 10 parents with children were asked to think aloud while using the app after a brief demonstration by the research assistant; in the second one, researchers included 32 parents in the study and the participants left the visit after the demonstration and used the app on their own for two weeks (Xiao et al, 2021). Post-study surveys with a Likert scale were also carried out to address the quality of instruction support, the ease of learning to take intraoral photographs and overall satisfaction (Xiao et al, 2021). Semi-structured interviews were conducted with the parents and the recordings were analyzed based on thematic content (Xiao et al, 2021). The System Usability Scale (Jordan et al, 1996) was also used to measure and quantify the perception of the product (Xiao et al, 2021).

Another study aiming to elucidate both the reliability and usability of the app Dental Trauma Tracker which was developed for the epidemiologic surveillance of traumatic dental injuries and the app's main stakeholders; health professionals, oral health professionals, and laypeople were included (Zaror et al, 2018). The system consists of two components, the app designed for general users who report dental trauma and the Website used by researchers to collect those reports to generate epidemiological data (Zaror et al, 2018). In addition to the validation of the app as a surveillance tool, the study comprised two design evaluations: an expert evaluation using Nielsen's (1994) "10 Usability Heuristics" and a testing of user interfaces with user participation (Zaror et al, 2018). In the user test, the participants were asked to complete 3 fictitious dental trauma cases, and three parameters were investigated: effectiveness measured by completion rate, efficiency measured by the time the participants took to complete a task, and satisfaction measured by the System Usability Scale (Jordan et al, 1996) and the Likert scale (Zaror et al, 2018).

Nielsen's (1994) heuristics and think-aloud method were employed in a study aimed at developing and testing the app ToothSense as an oral health promotion tool for the prevention of early childhood caries (Nolen et al, 2018). In the study, the participants were expected to complete the tasks while thinking aloud and to evaluate the app based on the principles of heuristics using the Likert scale (Nolen et al, 2018). The feature that distinguishes this study from other studies adopting Nielsen's methodology is that the people who evaluate using heuristics were the participant users, not experts.

A comprehensive design approach has been adopted in each of the development stages of the app aiming to facilitate the information exchange about dental traumas between patients and dentists (Stein et al, 2016). The design process was carried out in four stages: First, the researchers interviewed two dentists who worked in an emergency clinic to identify their needs and the type of information they review when a patient seeks urgent care (Stein et al, 2016). Secondly, they used the paper prototyping method to create a user flow based on the needs analysis in the first stage and paid attention to writing questions at an elementary reading level (Stein et al, 2016). In the third stage, the high-fidelity prototype was evaluated by the developers applying the "cognitive walkthrough" to determine how a novice user navigates through the application and by three experts through Nielsen's (1994) heuristics principles (Stein et al, 2016). In the final stage, after a timed-task completion assessment was made by including 20 participants, they were asked to think aloud retrospectively, and a survey was carried out to assess user satisfaction (Stein et al, 2016).

Another study presented the development process of the app aimed at informing patients in a written and visual form about implant treatments, which adopted a user-centered design approach and involved the user at almost every stage (Canbazoglu et al, 2016). In the ideation phase, developers/researchers surveyed 1610 dentists through an official dentistry web portal (www.dis-hekimi.org) to investigate their tablet ownership and opinions about the must-have content in the application, which the implant dentistry was chosen by 51% of the participants (Canbazoglu et al, 2016). After the user validation, the researchers used ethnographical observations and contextual inquiry methods to better understand the worksite environment without interacting with any staff, and then, they interviewed dentists to identify the treatment procedures to be included (Canbazoglu et al, 2016). In the following step, a workshop was organized including 11 dentists who were encouraged to discuss the findings of the previous stages and asked to sketch user interfaces in a limited time, so the participants could focus on the idea rather than the details (Canbazoglu et al, 2016). The user interface layouts were detailed with visual components and colors by using the paper prototyping technique and the ideas generated by the participants were criticized by the designers (Canbazoglu et al, 2016). After designing wireframes and high-fidelity interfaces and the coding phase, they included 100 dentists to test the app and asked them to complete the survey that aimed to collect demographic information, perceived ease-of-use, and perceived usefulness (Davis, 1989).

The assessment process of the app My Smile Buddy which was designed for community health workers who have no dental training and engage parents from low-income families was elucidated in a study (Chinn et al, 2013). The app aims to assist frontline health workers in assessing the caries risk in children, improving the oral health literacy of disadvantaged parents, and setting individualized goals for specific needs (Chinn et al, 2013). During the assessment and development phases of the app, five focus groups were conducted with the representatives of the local community of Northern Manhattan such as community health workers serving mothers with young children, staff from local organizations, and pediatric dental providers (Chinn et al, 2013).

As a quantitative measure, many studies employed different types of surveys to test usability and user satisfaction or collect sociodemographic data. After the apps were experienced with or without a given task, questionnaires were applied to children and their guardians for an app addressing dental caries, diet, and oral hygiene (Campos et al, 2019), to patients, and dentists for an app arranging dental appointments (Lin et al, 2014), to prospective users for an AI-powered app detecting caries (Liang et al, 2020), to the actual users of Brush DJ app (Underwood et al, 2015), to dentists for an app designed to collect patients' data (den Boer et al, 2018), to parents of 4 to 6-year-old children for an app addressing diet and oral health (Muhamat et al, 2021).

2.2.1.5 The issues of approaches in the mDentistry literature

Issues related to current articles can be addressed under two headings: the issues related to clinical assessments and design assessments by healthcare professionals.

Clinical approaches face several limitations. By its nature, other factors that may affect the outcome need to be standardized in order to obtain significant results on the benefits of those applications. For example, in a human trial to evaluate an oral hygiene app, factors related to the participants, such as tooth brushing habit, frequency, and effectiveness, must be at the same level so that the true effects of using such an app can be demonstrated. Therefore, in a controlled clinical study investigating the effect of an oral health app on oral hygiene, participants in both the test (the group consisting of participants who will use the application) and the control group (the group consisting of participants who will not use the application) undergo an oral hygiene training. However, a point that should not be missed is the possibility that people who already know that their teeth should be brushed twice a day, for at least 2 minutes, with the right technique and with fluoridated toothpaste, may not need to use these applications. Oral health apps can be beneficial for people who do not have access to such training provided in clinical trials and for those who do not have the awareness. There is a dilemma here. On one hand, clinical researchers need to eliminate or minimize inter-participant differences due to the aim of determining the real effectiveness of an app and the ethical reasons. On the other hand, when they eliminate or minimize these differences in order to provide ideal conditions in the study, they indirectly eliminate or minimize the chance of detecting the possible benefits that applications can provide in reality. As mentioned earlier, there are social determinants that affect oral health (AAPD, 2020). Particularly in underdeveloped and developing countries, people with low socio-economic status have limited access to oral health services, which may lead to an increased likelihood of developing oral diseases (AAPD, 2020; Jamieson et al, 2016; Antunes et al, 2006), and it is thought that oral health apps can alleviate this inequality (WHO, 2020). This dilemma arises from the discrepancy between the need to create ideal conditions for clinical trials and the fact that these conditions do not exist in reality.

A second factor that researchers should consider in their studies for oral health apps is the Hawthorne effect. This effect can be defined as a type of reaction in which participants change their behaviors in response to their awareness of being observed during the research (Roethlisberger & Dickson, 1939). Sedgwick and Greenwood (2015) stated that "the Hawthorne effect has considerable implications for the generalizability of results from research to clinical practice" (p. 2). This debate both in the medical and dental fields has prompted researchers to investigate the impact of this phenomenon (McCarney et al, 2007; Abdulraheem and Bondemark, 2018; Berthelot et al, 2019). It is worthy of note that participants included in oral health research may brush their teeth as they do not normally, and use the applications provided to them only because they are part of the research.

Another recently discussed shortcoming in controlled clinical trials investigating the potential efficacy of digital health products is the non-use of a "*digital placebo*" (Torous et al, 2022). The concept was first introduced by Torous and Firth (2016). The digital placebo can be defined as the effect of using only a digital product, apart from the “*active ingredient*” that has a potential effect, in interventions made through mobile applications. At this point, it is necessary to investigate how important this effect is in measuring the effectiveness of oral health applications, and if there is a significant effect, it is also necessary to design functionless placebo applications without the “*active ingredient*”.

There are two main shortcomings in the studies that evaluate the design and usability of oral health apps. Healthcare professionals (Tiffany et al, 2018) evaluated 33 oral health applications using content analysis and a method called "Persuasive Health Technology Heuristics" (Kientz et al, 2010) inspired by Neilson's (1994) "10 Usability Heuristics" from the design literature. The fifth item of “Persuasive Health Technology Heuristics” (Kientz et al, 2010) is related to the usability and aesthetic appearance of the evaluated product. However, Tiffany et al. (2018) evaluated those features of oral health applications without involving the end-user. The human-centered design approach underlines the necessity of putting the user's needs and interests at the center of the process while the products are still in the design phase. Evaluating usability without involving end users means excluding them completely from the process. Because these people may not have been interviewed or even asked for their opinions during the design phase. As mentioned before, it is only possible to determine whether those apps are truly usable by meeting the user with the product and observing how the interaction occurs. In the absence of this, it is not the right approach to draw a conclusion, and healthcare professionals should avoid such approaches.

The second fallacy of the existing studies aiming to determine the usability of oral health apps is that they only adopt quantitative methods and draw conclusions based on the averages. Quantitative methods are frequently used in the fields of medicine and dentistry, and the researchers from these fields having such reflexes can be understandable. When the user’s culture, lifestyle, digital literacy, or competence come into play, such methods are not sufficient to collect the necessary data for usability. For instance, does an average score of 3.5 as a result of a usability survey truly indicate that the product is usable? In quantitative research, it is difficult to understand the motivations of some users who give low scores. The valuable data that could have actually guided a designer or developer in improving the product’s usability gets lost among the numbers in such studies. For this reason, qualitative methods such as in-depth interviews or think-aloud rather than quantitative methods will undoubtedly have greater contributions to studies on usability.

In conclusion, the field of mDentistry is a fairly new field and various challenges await researchers working in this field. New methods and standards are needed to obtain reliable results in studies to determine their clinical efficacy. If physicians have legal responsibility after recommending health apps, researchers should tackle those products more comprehensively, including diverse user profiles and avoiding generalizing approaches. It shall not be forgotten that apps cannot be treated like medicines and their efficacy cannot be tested in laboratory conditions. As mentioned earlier, there are many factors affecting user-app interaction. Therefore, studies of digital oral health products should be carried out with a multidisciplinary approach in collaboration with different fields such as design, dentistry, health psychology, and sociology.

2.2.2 Apps as medical devices

According to a report (Statista, 2022), there are approximately 55,000 mHealth apps (excluding well-being and fitness apps) available on the Google Play Store in 2022, and after a decline in 2018, they have increased again with the COVID-19 pandemic. Throughout the pandemic, the challenges posed by mandatory quarantines have shown us how much we need more digital health solutions to overcome restrictions on access to healthcare and information. Health apps have since been used more and more to diagnose, treat, and manage diseases, complementing existing healthcare, and investors realizing this increased their investments in these technologies (Kadakia et al, 2020). Based on the potential of mHealth and the burden of oral health diseases on societies, the WHO (2021) made the development of the "mOralHealth" program one of its priorities and published a guideline to contribute to the achievement of better oral health through digital tools. However, due to raising concerns about the lack of evidence for their efficacies, low-quality health content, false claims, and privacy risks in studies (e.g., Tiffany et al, 2018; Chen et al, 2021; Martinengo et al, 2019; Huckvale et al, 2015), the authorities stepped in and imposed more strict regulations to improve the standard of apps, especially for those that involve diagnosis, treatment, and management of chronic diseases or high-risk conditions.

2.2.2.1 Governing regulations for medical apps

Digital products meeting the definition of software as a medical device (SaMD) are defined as "*software intended to be used for one or more medical purposes without being part of a hardware medical device*" (International Medical Device Regulators Forum (IMDRF) SaMD Working Group, 2013, p. 6). The US Food and Drug Administration (FDA) is responsible for approvals and regulations of SaMD like all medical devices in the United States and issued the Policy for Device Software Functions and Mobile Medical Applications in 2013 and subsequently updated in 2015, 2019, and 2022 to accelerate the pace of digital innovation and inform manufacturers, distributors, and other entities about how the agency intends to apply its regulatory legislation (U.S. Food and Drug Administration, 2022). It applies its regulatory oversight to only those software functioning like medical devices that could pose a risk to public safety if they were not to function as described by the manufacturer (U.S. Food and Drug Administration, 2022). The requirements to be met by the manufacturer vary according to the level of risk posed by the medical app, and thereby they are correspondingly classified under three classes: class I, II, and III medical devices. Class I devices require general controls, while class II require special controls in addition to general controls and class III must be approved by the authority before market entry (U.S. Food and Drug Administration, 2022, p. 11).

FDA considers the following software functions as devices being subject to regulatory oversight:

1- "*Software functions that are an extension of one or more medical devices by connecting to such device(s) for purposes of controlling the device(s) or analyzing medical device data*" (U.S. Food and Drug Administration, 2022, p. 11).

An example of these types of functions includes apps that enable users to control the delivery of insulin on an insulin pump by transmitting signals (U.S. Food and Drug Administration, 2022, p. 11).

2- "*Software functions (typically, mobile apps) that transform the mobile platform into a regulated medical device by using attachments, display screens, or sensors or by including functionalities similar to those of* *currently regulated medical devices. Software functions that use attachments, display screens, sensors, or other such similar components to transform a mobile platform into a regulated medical device are required to comply with the device classification associated with the transformed platform*." (p. 12)

An example of such functions is applications that work in coordination with blood sugar measuring strip readers connected to the mobile platform and that display the results on the screen (U.S. Food and Drug Administration, 2022, p. 12).

3- "*Software functions that become a regulated medical device by performing patient-specific analysis and providing specific output(s) or directive(s) to health care professionals for use in the diagnosis, treatment, mitigation, cure, or prevention of a disease or condition. Additionally, software functions that perform patient-specific analysis and provide patient-specific diagnosis or treatment recommendations to patients, caregivers, or other users who are not health care professionals." (p. 12)*

An example of such functions is software that analyzes the patient's medical data to detect a life-threatening condition, such as a stroke and generates an alarm to notify the user and his/her physician (U.S. Food and Drug Administration, 2022, p.12), (for further information about other software functions that are subject to regulatory oversight, see U.S. Food and Drug Administration, 2022, Appendix C).

For some software in the following examples, FDA intends to exercise enforcement discretion:

1- "*Software functions that provide or facilitate supplemental clinical care, by coaching or prompting, to help patients manage their health in their daily environment*" (p. 14).

These are mobile apps that assist patients with chronic diseases such as hypertension, diabetes, or obesity to maintain a healthy weight, stay fit, reduce salt intake, or remember to take medicine by simple prompting (U.S. Food and Drug Administration, 2022, p.14).

2- “*Software functions that are specifically marketed to help patients communicate with health care professionals by supplementing or augmenting the data or information by capturing an image for patients to convey to their health care professionals about potential medical conditions.*” (p. 14)

An example of such software functions is the telehealth apps that connect healthcare professionals to patients through a camera to transmit images and enable consultation remotely (U.S. Food and Drug Administration, 2022, p.14).

3- “*Software functions that perform simple calculations routinely used in clinical practice*” (p. 15).

These apps facilitate practitioners to perform simple and routinely used calculations in medical practice such as Body Mass Index, Total Body Water/ Urea Volume of Distribution, or delivery date estimator (U.S. Food and Drug Administration, 2022, p.15).

It is worthy of note that the apps aiming at reminding or tracking oral hygiene practice are examples of software functions that may meet the definition of medical device and the FDA, therefore, intends to exercise enforcement discretion (for further information see U.S. Food and Drug Administration, 2022, Appendix B).

The software functions described below are not considered medical devices by the FDA, so they are not subject to regulatory requirements:

1- “*Software functions that are intended to provide access to electronic “copies” (e.g., e-books, audiobooks) of medical textbooks or other reference materials with generic text search capabilities*” (p. 17).

These are the apps that include medical dictionaries, electronic copies of medical textbooks, or diagnostic manuals (U.S. Food and Drug Administration, 2022, p. 17).

2- “*Software functions that are intended for health care professionals to use as educational tools for medical training or to reinforce training previously received*”(p. 18).

These are the apps that include medical flashcards with images and graphs, interactive anatomy diagrams, surgical training videos, or preparation content for medical board certification (U.S. Food and Drug Administration, 2022, p. 18).

3- "*Software functions that are intended for general patient education and facilitate patient access to commonly used reference information*" (p. 18).

Examples of such apps include educational information for patients regarding their disease, condition, treatment, or upcoming procedure, lists of emergency hotlines, physician advice lines, or locations of nearby pharmacies (U.S. Food and Drug Administration, 2022, p. 18).

4- "*Software functions that automate general office operations in a healthcare setting*" (p. 19).

Examples include software functions that determine billing codes, enable insurance to collect and process data, perform medical business accounting functions, schedule medical appointments, or track patient satisfaction (U.S. Food and Drug Administration, 2022, p. 19).

5- "*Software functions that are generic aids or general-purpose products*" (p. 19).

Examples of these apps function as recording audio, taking notes, or facilitating communication between healthcare professionals in a hospital (U.S. Food and Drug Administration, 2022, p.19).

6- "*Software functions that are intended for individuals to log, record, track, evaluate, or make decisions or behavioral suggestions related to developing or maintaining general fitness, health or wellness*" (p. 19).

These are apps that include information about healthy diet, exercise and weight loss, calorie counters, or track a baby's sleeping and feeding habits (U.S. Food and Drug Administration, 2022, p. 19). The other functions described in the rest of the articles in Appendix A are generally related to those that enable patients or healthcare providers to record health data in an online database (for further information see U.S. Food and Drug Administration, 2022, p.20-23).

Prior to the entry into the European market, medical devices must have a CE mark and have been regulated through EU directives called Medical Device Directive (MDD) and Active Implantable Medical Device Directive (AIMDD) for a long time (OJ L 169, 12.7.1993., as cited in Malvehy et al., 2021). In 2017, after the scandals involving accidents due to unsafe medical devices, the European Commission decided to revise the MDD 93/42/EEC (OJ L 169, 12.7.1993.) and replaced it with the Medical Device Regulation (MDR(EU)) 2017/745 (OJ L 117, 5.5.2017.) (as cited in Holfelder et al., 2021) by keeping its core elements (Niemiec, 2022). This new regulation which came into effect on 26 May 2021 and is enforceable as law in all member states unlike MDD tightens controls on medical devices and requires the manufacturers to prove their performance and safety. The risks are expected to be acceptable and the benefits to pay off. In MDR, medical devices are defined under four categories taking into account their intended purpose and inherent risks: class I (lowest risk), class IIa, class IIb, and class III (highest risk) (OJ L 117, 5.5.2017., see Article 51 and Annex VIII).

Although it does not give much detail, the classification rule related to software in MDR was described as follows (OJ L 117, 5.5.2017., see Annex VIII / Chapter III):

> "*Rule 11*
>
> *Software intended to provide information which is used to take decisions with diagnosis or therapeutic purposes is classified as class IIa, except if such decisions have an impact that may cause:*
>
> > *— death or an irreversible deterioration of a person's state of health, in which case it is in class III; or*
> >
> > *— a serious deterioration of a person's state of health or a surgical intervention, in which case it is classified as class IIb.*
>
> *Software intended to monitor physiological processes is classified as class IIa, except if it is intended for monitoring of vital physiological parameters, where the nature of variations of those parameters is such that it could result in immediate danger to the patient, in which case it is classified as class IIb.*
>
> *All other software is classified as class I."* (p. 188)

It is worthy of note that not all software used within healthcare is considered a medical device such as those with a library, invoicing, or staff planning functions, however, the software intended to process, analyze, create, and modify medical information with or without influencing the use of hardware medical device may be qualified as a medical device and fall under the MDR (Medical Device Coordination Group, 2019, p. 6). As aforementioned in "*Rule 11*", the risk of harm to patients, users of the software, or any other person related to the use of the software plays a central role in qualifying an app as a medical device. Therefore, for software entailing such risks, manufacturers must ensure that all regulatory requirements and conformity assessments have been fulfilled and some must be supported by clinical evidence to be Conformité Européenne (CE) marked, which is mandatory for placing on the European market (Medical Device Coordination Group, 2019, p. 8). According to "*Rule 3.3 of Annex VIII*" (OJ L 117, 5.5.2017), the software driving or influencing the use of a medical device falls within the same class as the device (hardware), while software that is independent of any other device is classified in its own right. In the guide (Medical Device Coordination Group, 2019, p.27) the classes were exemplified as follows:

- Software aiming to perform diagnosis through image analysis for making treatment decisions in patients with acute stroke is classified as class III.

- A mobile app intended to analyze a user's heartbeat, detect abnormalities, and inform a physician accordingly for guidance in diagnosis is classified as class IIb.

- Software intended to be used to monitor vital physiological signals in routine check-ups or at home is classified as class IIa.

- A mobile app aiming to support conception by calculating the user's fertility status is classified as class I.

Independent of the risk class, a manufacturer must compile technical documentation for the medical app to declare that it meets the MDR requirements for general safety and performance (OJ L 117, 5.5.2017., see Annex II). Annex II lists the required documentation including risk management system, benefit-risk analysis, clinical evaluation report, software life cycle file, usability file, and others (OJ L 117, 5.5.2017., see Annex II, as cited in Keutzer & Simonsson, 2020). Except for class I apps, all other classes must be assessed and approved by a notified body which is an organization designated by national authorities to check the devices' compliance with standards and rules specified by MDR before being placed on the market (Keutzer & Simonsson, 2020). In addition to a notified body, class III apps must be assessed by an expert panel as to whether the clinical evidence is sufficient to provide confidence for its safety and performance (OJ L 117, 5.5.2017, see Annex IX). Conformity assessment pathways, required certificates, standards, and documents for medical apps vary from case to case and are very detailed, which goes beyond the aim of this thesis. (For introductory information about the process, read Keutzer & Simonsson, 2020; Holfelder et al, 2021).

Other binding regulations for medical apps are the General Data Protection Regulation (GDPR) (European Parliament, 2016) in the European Market and the "*Health Insurance Portability and Accountability Act of 1996*" (HIPAA) in the United States. The manufacturer needs to ensure that the app does not encroach upon patients' privacy and protects their health data. The principles relating to the processing of personal data in GDPR (Article 5) state that the data must be processed lawfully, fairly, and in a transparent manner; collected for specified, explicit, and legitimate purposes; adequate, relevant, and limited to what is necessary concerning the purposes; accurate and kept up to date; kept in a form which permits identification of data subjects for no longer than is necessary; protected against unauthorized or unlawful processing and accidental loss or destruction; and that the controller shall be responsible for the compliance. Briefly, the manufacturer shall ask permission to use data, declare the purpose of its use, and with whom it will be shared. Apart from the regulations for data protection, in order to track each device, react quickly in case of serious incidents, and prevent placing medical devices without compliance on the market, both FDA (2020) and EU (OJ L 117, 5.5.2017, Article 27 and Annex VI) developed a system called the Unique Device Identifier (UDI) to link the device to their database, so that all relevant information will be accessible (as cited in Keutzer & Simonsson, 2020).

Whilst the number of AI-powered medical software approved by authorities in the USA and Europe has increased in the past years (Muehlematter et al, 2021), the systematic reviews indicate that there is a lack of standards, high-risk bias, heterogeneity in methodologies, and poor reporting in the existing studies showing the diagnostic performance of these devices (Liu et al, 2019; Aggarwal et al, 2021; Nagendran et al, 2020). Because of these drawbacks, there may be an overestimation in these studies showing the performance and safety of medical AI systems, which entails a great risk of harm to patients if the devices are used in practice (Niemiec, 2022). Like other medical apps, AI systems developed for medical care also fall under the regulations. However, machine learning (ML) technologies are more difficult to audit due to their ever-changing nature. Regarding this issue, the FDA (2021) has released the Artificial Intelligence as a Medical Device/Machine Learning (AI/ML) Based Software (SaMD) Action Plan. This document summarizes the feedback received from the stakeholders and describes a five-part Action Plan. In the first part, the agency undertakes to develop a tailored regulatory framework for AI/ML-based SaMD. The framework aims to guide manufacturers to describe "*what*" aspects they intend to change through learning and "*how*" the algorithm will learn and change while remaining safe and effective (U.S. Food and Drug Administration, 2021, p. 3). In the second part, the agency acknowledges the need for describing the "*Good Machine Learning Practice*" not only to provide guidance in the industry but also to facilitate oversight of these devices and highlights its commitment to deepening its participation in AI and Medical Device communities to encourage consensus (U.S. Food and Drug Administration, 2021, p. 4). In the third part, it is encouraged to adopt a patient-centered approach during the development process to improve usability, and users' trust, and emphasizes the importance of transparency, which is needed for manufacturers to ensure users understand how AI functions and what kind of benefits, risks, and limitations they have (U.S. Food and Drug Administration, 2021, p. 5). In the fourth part, it is mentioned that the data collected from historical datasets to train algorithms are vulnerable to bias due to the inequality in access to healthcare, so these algorithms must be well-suited for people of different races, ethnicities, and socioeconomic status (U.S. Food and Drug Administration, 2021, p. 5). In the final part, it is promoted to collect data on the real-world use of the devices to understand how the product is used and identify the opportunities for design refinements, and usability concerns (U.S. Food and Drug Administration, 2021, p. 6).

2.2.2.2 The role of designers in the medical device development processes

Due to complex development and long pre-market launch processes under heavy regulations, the medical device industry has its own unique challenges. Medical devices have many stakeholders that play key roles in their journey, and these stakeholders are not limited only to patients using these devices but also patients' caregivers, manufacturers, insurance companies, regulatory agencies, healthcare professionals, hospitals, and medical device providers (Privitera, 2017). Designing a medical device is more complex than designing a consumer product, and traditional definitions of design-driven innovation (Verganti, 2009) with the consumer product in mind may remain an overly simplistic definition for the field (Dunn et al, 2019). The risk of technical failure of a medical device is more critical and vital than that of a daily consumer product. For this reason, the design processes need to be carried out meticulously. Only when certain quality standards are reached and the necessary certifications are obtained, these products can enter the market. Because of all these challenges, medical device processes are different from other products. Developing a successful medical device requires both engineering and clinical, regulatory, marketing, and business expertise (Panescu, 2009).

Industrial design is a professional service aimed at creating and developing concepts to optimize the functionality, value, and appearance of products and systems that are mutually beneficial to both the user and the manufacturer (Goldberg et al, 2007). Ulrich and Eppinger (2008) identify industrial designers' main responsibilities as designing user-related aspects of the product such as aesthetic and human factors, they, therefore, can play a vital role in ensuring that the technical team does not overlook user needs. In the design process, the designer plays a major role in important areas such as creativity, as well as being a mediator for the interests of different parties and a facilitator for the ideas of the participants (Taylor et al, 2011). Margolis and Pauwels (2011) state that as an interdisciplinary field, design is fed from different fields such as craft, science, creativity, commerce, humanities, social studies, art, and engineering. Considering all these skills, the contribution of a designer to the medical device design process is quite meaningful.

Technology-related aspects of a product are often kept separate from user-related aspects (Jones, 2013). However, the user-involving design approach should be adopted in all medical devices by the manufacturers to ensure that the needs of the user are met as much as possible and to prevent possible use-related errors through design (Schaeffer, 2012) by increasing the usability of the device. Although in recent years, efforts for usability in medical devices have ceased to be an arbitrary action and have become a mandatory process through regulations for manufacturers, no comprehensive framework for user involvement exists. (Privitera et al, 2015). Regarding this, the manufacturers' motivation to conduct user research cannot go beyond fulfilling the regulatory compliance process (Martin & Barnet, 2012) and the user involvement usually takes place after the form and functions of the device have already been determined (Throckmorton, 2016).

Usability tests provide important insights into product interfaces and how they communicate with users. The user's participation after the product is completed to tick the box in the list of requirements increases the risk of experiencing problems during the interaction with the product and encountering unexpected negative events. In cases where adverse events are encountered, the device manufacturers often take the position that it is sufficient to explain "*correct operation*" in the instruction manual and that the problems experienced while using the device are solely related to the user (Fairbanks & Wears, 2008). However, Fairbank and Caplan (2004) have identified many problems in the interfaces of medical devices on the market and noted that many adverse events in the healthcare field are caused by problematic design rather than user error. Poor design cause not only a decrease in their motivation for the treatment and use of the device (Fung et al, 2015a; Fung et al, 2015b) but may also result in critical sanctions for the manufacturer such as the withdrawal of the product from the market (Harte et al, 2014). For these reasons, it is a necessity to design the interaction by considering the ergonomics principles in the development processes and to test the usability of the devices by including the users at every stage of the product.

Recent studies (Privitera, 2017; Köse & Cifter, 2018) conducted with medical device manufacturers reveal the perception of the role of the designer in this sector. In interviews with 18 manufacturers from the USA and EU, Privitera (2017) reported that 12 of them do not employ industrial designers, the remaining 6 companies employ in-house designers and 5 of them receive design services from consulting companies. When these manufacturers were asked about the role of the designers in medical device development, 13 companies stated that the primary role of the designer was to make the device more aesthetic, 9 companies described their responsibility as human factors, 6 companies noted other functions such as the concept design and determining the needs, and 4 companies reported the role of branding (Privitera, 2017). The study (Köse & Cifter, 2018) that included 17 medical device manufacturers from Türkiye showed that only 24% of the respondents employ at least one designer, and 41% opt for work with a consultancy firm. The participants were asked what the 6 requirements mentioned in the human-centered design literature mean for their processes to measure their awareness (Köse & Cifter, 2018). It was reported that the usability of the medical device, user interface, use error, user experience, ergonomics, and patient safety were mentioned by 40%, 67%, 80%, 67%, 47%, and 87% of the respondents respectively (Köse & Cifter, 2018).

The respondents described their experiences with industrial design which led the author to categorize the challenges for designers (Privitera, 2017). The biggest challenge is that the designers' contribution to the process in return for their high costs cannot be fully defined (Privitera, 2017). The second important challenge is that the designers do not have sufficient technical and medical knowledge about the field, causing them to not understand the process and consequently their contribution to be limited to aesthetic design (Privitera, 2017). Another challenge is that manufacturers' view of the human factor is often something that needs to be done for regulations (Privitera, 2017).

2.2.2.3 Prospects and challenges

As mentioned earlier, medical apps offer great opportunities in healthcare. When making use of these products, it must be ensured that they do not pose risks to public health. In this context, regulations will minimize these risks by requiring that these devices have a certain standard in terms of quality and safety. Hence, this will facilitate their widespread adoption in the healthcare field and encourage the advancement of the mHealth industry in general. However, they can also lead to some challenges and pitfalls that manufacturers may encounter in the development and compliance processes if necessary precautions are not taken by the authorities.

In the literature, medical app-related issues such as lack of evidence for their efficacy, poor quality health content, false claims, and privacy risks have already been highlighted (Tiffany et al, 2018; Chen et al, 2021; Martinengo et al, 2019; Huckvale et al, 2015). In this sense, these strict regulations governing the apps will force manufacturers to comply with standards that will increase the quality of these devices in terms of medical content, data privacy, efficacy, and usability regardless of their risk classes. Regulations such as GDPR or HIPAA will improve their security and protect users' sensitive data against possible breaches. UDI systems will enable products to be traceable and in case of any adverse events and omissions, it will be possible to withdraw them from the market quickly and initiate the necessary legal actions against those responsible.

It's reasonable to say that, given the growing number of medical applications on the market, choosing one can be difficult. Physicians face the same struggle while recommending the most appropriate app for their patients and furthermore, they may be held legally responsible for any problems that may arise with the apps they recommend (Palmer, 2017). New regulations that impose more stringent requirements will provide a solution to this problem. Only apps whose compatibility is overseen by the authorities will be available on the regulated markets, resulting in a drop in numbers and an increase in quality. Hence, healthcare providers and patients/users will be able to rely on these technologies more than ever, which will eventually encourage their use in clinical and daily practice.

Since mobile applications are relatively new technologies, no consensus has yet been reached on how clinical and design studies should be carried out in this field. However, the study designs will also be indirectly affected by these developments brought about by regulations. Researchers will put forward new suggestions regarding the methods, such as using placebo apps (Torous et al, 2022), to make these assessments in a more standardized and bias-free manner. In this way, the data quality will increase, and this will help researchers to better understand the parameters that affect the performance of these devices.

While the accreditation process is decentralized in some countries such as Sweden, most of them have embraced a centralized approach, which ensures consistency in approvals but entails "*a risk of creating bottlenecks*" (Essén et al, 2022). Considering the large number of apps pending approval, it is an important issue that may particularly affect small-scale manufacturers such as start-ups. Delays and long periods in approval processes will increase the overall cost, which is already high for these companies. In the article, in which the application processes of two medical apps named "Corona Check" and "Corona Health" are elaborated, it is stated that the preparations for the harmonization process of the two apps cost a total of €410.000 (Holfelder et al, 2021). Since not every company in the industry can afford it, this may lead to a situation where only large-scale firms can survive in the market and atrophy of innovation potential, enthusiasm, and diversity of ideas. Rising costs will also aggravate the financial burden on the manufacturers, which will eventually affect end users. One of the prominent benefits of apps is that they offer a great opportunity to alleviate inequality in access to healthcare services and information. However, manufacturers may be forced to change their business models and price their apps in order to survive, which may hinder all segments of society from benefiting from this technology.

In the EU, the notified bodies that are authorized by the state's public health agency to evaluate the medical devices have the right to issue a CE mark, which is mandatory for manufacturers before marketing their products. There are currently 34 notified bodies in the EU (European Commission, n.d., Accessed in 2022). Notified bodies are private companies that provide conformity services for a certain fee to manufacturers who want to obtain the CE mark. It was indicated that manufacturers are customers of these bodies, so they compete from a pricing standpoint and may be reluctant to not approve devices to maintain their business relationship (Van Norman, 2016).

Policymakers should be aware of potential risks that could slow or hinder the compliance processes of medical apps. Legislation regarding these digital technologies should follow the progress in this field and descriptions should be determined and updated in a way that leaves no room for interpretation. To achieve that, agencies should be in close contact with communities of stakeholders by organizing workshops and proactively engaging in ongoing discussions. Based on the feedback and suggestions they receive; new guidelines should be published to prevent errors due to misunderstandings. The procedures should be designed to audit the competencies of these devices in safety and performance fairly, avoiding over-regulation. Small-scale manufacturers should not be left alone in approval processes and supported both technically and financially.

3. RESEARCH APPROACH

Most research in medicine adopts a traditional positivist principle, where researchers are expected to be impartial and objective. The knowledge is based on observations and measurable facts, and the inquiry should be done without the influence of personal biases. Since there is no room for researchers' subjective opinions and interpretations, quantitative methods, and statistical evaluations are often at the center. To be able to analyze the data quantitatively, researchers approach the subject in a reductionist way by separating complex phenomena into simpler components. They pay attention to controlling extraneous variables to minimize their influence on the results so that the relationship between the variables can be determined with confidence where the size of the "*sample*" often plays an important role. The knowledge is generated cumulatively; therefore, the methods and results should be repeatable and reproducible by other researchers.

However, Susman and Evered (1978) state that in organizational science, researchers use positivist methods, which have great value in natural sciences such as biology and physics, and even in social sciences, but these methods have become sophisticated and far from solving the problems faced by the members of the organization. The use of the positivist model in organizations in order to be neutral, logical, and empirical, on the other hand, causes the values of the members to be overlooked and the knowledge produced through research cannot address a solution to the current problems (Susman & Evered, 1978). This crisis results from the separation of theory from practice. Theories fail to meet the needs of practitioners, so research papers are read more by research producers than practitioners (Susman & Evered, 1978). Schon (1984, as cited in Meyer & Batehup, 1997) believes that positivism's "*technical rationality*" has shaped the way of research, education, and professional thinking, therefore, this practice of technicality, based on specialized scientific knowledge, leads to a hierarchical model which separates research from practice. He argues (Schon, 1984) that a theory-practice gap emerges from this separation, and that we witness this gap in the work life of professionals where they often rely on their intuition because sometimes the technical knowledge does not fit the situation, they are in. Schon (1984) identifies this kind of knowledge that derives from intuitive practice as "*knowing-in-action*".

Some authors criticize the positivist approach in several ways. First of all, they state that the perception that the methods used in the positivist approach are value-neutral is a misconception, that at the end of the day, the researcher determines these methods, and that it would not be wrong to describe the researcher as a manager rather than an experimenter (Susman & Evered, 1978). They noted that the positivist approach to science is itself a product of the human mind, therefore human consciousness, culture, views, and every factor in the social context that may influence the research process cannot be excluded (Susman & Evered, 1978). In this traditional paradigm, in which all human characteristics are ignored, people are treated as objects of inquiry (Susman & Evered, 1978), or subjects who provide data and respond in information retrieval processes (Reason & Bradbury, 2001, as cited in Hockley & Froggatt, 2006). For all these reasons, it has been stated that the positivist approach is not appropriate in social studies dealing with human inquiry (Meyer & Batehup, 1997).

Lewin (1946) points out the need for a new approach that researchers can adopt in their studies focused on practice by cooperating with practitioners in making a change. He identifies a framework for action research, emphasizing its participatory characteristics, democratic impulse, and contribution to social science. Later, Susman and Evered (1978) identify action research as a more "*congenial*" mode of inquiry and state that it can correct the deficiencies of the positivist approach in the social context with its "*future-oriented*", "*collaborative*", "*problem-solving*", "*agnostic*", and "*situational*" nature by "*generating theories grounded in action*". It is an act of theory development within a practice discipline, where action researchers do research "*with*" or "*for*" people rather than "*on*" people by gaining an inside view of their experience and shedding light on tacit knowledge of human situations (Meyer & Batehup, 1997). In a social context where variables cannot be controlled, the researcher may feel trapped by the strict standards of the positivist methodology. Action research is an emancipating approach that enables researchers to deal with ever-changing circumstances where the conditions are unpredictable and unfold as the research progresses (East & Robinson, 1994). In the presence of these unpredictable conditions, this approach focuses on the process itself as well as the outcomes of the innovation, therefore, it allows the process of change to be explored rather than simply intended to measure output (Meyer & Batehup, 1997). Baskerville (1999) underscores that introducing changes in these processes and observing their effects is the best way to study complex social systems because such systems cannot be simplified easily into meaningful studies. Therefore, researchers should investigate human organizations, particularly those interacting with information technologies, as "*whole entities*" (Baskerville, 1999). The researcher should enter into and be part of this complexity where the process, the ongoing change, human agencies, and everything related to these factors intertwine to better understand the situation. Here, the success of the researcher is not only the extent of change achieved or how the solution is brought about but also what has been learned from the experience (Meyer, 2000).

Meyer and Batehup (1997) note that it is difficult to define what action research is and that therefore, it should be considered an approach rather than a method. Although four types of action research (experimental, organizational, professionalizing, and empowering) have been identified (Hart & Bond, 1995), its "*multidimensional nature*" rarifies classifying individual studies (Meyer, 2000). The reason why action research is used "*loosely*" in the literature may be that the approach is shaped according to the unique situation that each researcher is in. It changes in form and responds to specific problem areas. In fact, Meyer and Batehup (1997) state that researchers[5] should develop their own action research forms by being eclectic and pointing out the weaknesses and strengths specific to their practical situation so that only in this way researchers can bridge the gap between theory and practice. Greenwood (1994) emphasizes the importance of critical thinking and methodological debate when researchers transfer a method from one discipline to another. She argues that researchers in the health field should be aware that there may be some implications of such "*emancipatory action*" and that what might be appropriate for the educational settings might not be suitable for healthcare practice (Greenwood, 1994). In other words, researchers in the field of health, who

[5] The authors point to nurses because their article is on action research in nursing.

will carry out action research, can develop their own method according to their specific situations, but they should not forget to discuss their methodology critically.

Literature has different thoughts about what action research should focus on. The gap between theory and practice has been criticized because the theory produced in research with a positivist approach does not come up with a solution to the problems of practitioners. When Susman and Evered (1978) state that action research is "*grounded in action*", they refer to the grounded theory which is described as appropriate in the study of social interaction or experience where researchers try to explore a situation, not test, or validate an existing theory (Lingard et al, 2008). For the experience and results to be generalizable and to rebuild the bridge between theory and practice, some authors underscore that action research needs to generate a theory in light of the data gathered from practice (Susman & Evered, 1978; Meyer & Batehup, 1997). However, it is thought that action research may miss the point when researchers prioritize academic concerns rather than the reason for doing the research. Therefore, Elliott (1991) argues that action researchers should focus on diagnosing the problem and improving the practice rather than generating theories.

Various criticisms and warnings have been made for action research. Most of these criticisms are on the validity of the knowledge obtained through action research. From the natural science view, action research lacks the characteristics of scientific inquiry where the knowledge gathered can be reduced, repeated, and refuted. While researchers try to understand what is observed, they draw a conclusion according to their understanding, which leads the researchers' perspectives to be reflected in all observation records (Kant, 1908). Since all conditions and parameters are standardized in studies of the natural sciences and thus can be repeated by other researchers, a bias, if exists, will eventually come into view. However, action research is situation-specific, where the same conditions cannot be reproduced, and the research cannot be repeated. In action research, the researcher intervenes in the ever-changing process, thus becoming a part of the study and one of the study subjects. Therefore, interpretative statements, the observer's values, and prior knowledge intrude upon the observation (Baskerville, 1999).

Action research must be evaluated according to a completely different set of criteria, and generalizations through action research have different dynamics from empirical research (Meyer & Batehup, 1997). Susman and Evered (1978) argue that if being scientific requires only a positivist approach, then action research is definitely not scientific but a different kind of inquiry that contributes to the growth of knowledge, therefore it is legitimate. Meyer and Batehup (1997) point out that there may be a risk of lack of clarity about what data collected during action research is used, analyzed, and interpreted to contribute to knowledge and that transparency of data is imperial, especially in the health context. Therefore, Simons (1981) underscores that action researchers should pay attention to writing detailed, rigorous, impartial, and accurate reports that are accessible and evaluable by both the public and professionals. To avoid such risks, there should be a section in the study, called self-reflective field notes, in which the researcher clearly states his or her own beliefs, values, and perspective so that the reader can evaluate the research more clearly by understanding from which perspective the researcher approached the situation.

For the last three decades, action research has also gained visibility in medicine because there was an increasing concern that practice was not always sufficiently influenced by evidence (East & Robinson, 1994). Action research has been utilized in studies, where researchers needed to adapt to changing conditions in health settings as in the example of a nursing home for dying older people (Hockley & Froggatt, 2006). To the best of our knowledge, there is no action research exploring the design process of a medical device in the literature. Although the human relations in the health environment and the design of a health product have different dynamics, the need for adapting to the constant change of the process is similar.

Action research and design research have many similarities in terms of approaching a subject. Järvinen (2007) analyzes both approach conceptually and defines seven fundamental similarities: 1) emphasis on the value of utility, 2) production of knowledge to guide practice, 3) combination of taking action and evaluation, 4) collaboration of research and practice, 5) aiming for development and improvement, 6) intervention to a problem and 7) knowledge is generated, used, and evaluated through the process. Similarly, Cole et al (2005) state that both research approaches share common features in terms of ontology, epistemology, and axiology. Ontologically, the phenomenon of interest does not remain static throughout the research process in both approaches (Cole et al, 2005). Epistemologically, they adopt a mode of knowing that involves intervening to effect change, and reflecting on this intervention (Cole et al, 2005). Finally, they axiologically emphasize practical utility and theoretical knowledge (Cole et al, 2005).

3.1 Research in Design

Frayling (1994) classified the research conducted in the field of art and design under three headings: "*Research Through Design*" (RtD), "*Research for Design*" (RfD), and "*Research into Design*" (RiD).

RtD is a type of research in which design practice is a part of the methodology, taking the process into the center, and where the research and design are intertwined. Therefore, the research process is as iterative as the design process. Zimmerman (2010, p. 310) states that it offers distinct advantages to researchers allowing them to rely on the design activities when approaching "*messy situations with unclear or even conflicting agendas*" where other inquiry methods are not suitable. Frayling (1994) finds it the least straightforward because it lacks structures and models from which to derive its rules and procedures like other types but is still an identifiable and visible type of research. Researchers adopting RtD create new products, experiment with new materials and processes (Findeli, 2004, as cited in Godin and Zahedi, 2014), and document the results of the design practice rigorously and explicitly to produce communicable knowledge. In this sense, it resembles action research in many ways. It is difficult to adopt such a method for a researcher who is not experienced in design practice (Megahed, 2017). Because, just as the researcher turns into a designer or vice versa in the process, the process itself turns into the material of the research. The designer or researcher becomes not just an observer but also an actor and a participant observing his/her own creative process. Glanville (1997, as cited in Megahed, 2017) describes RtD as research where the researcher is inside looking outward.

RfD involves investigations conducted with an idea in mind and actions to find a way to bring it into being, in Fraser's words, it enables the design to occur (Fraser, 2013 as cited in Megahed, 2017). Frayling (1994) describes RfD as:

> *"...Research where the end product is an artefact -where the thinking is, so to speak, embodied in the artefact, where the goal is not primarily*

communicable knowledge in the sense of verbal communication but in the sense of visual or iconic or imagistic communication." (p. 5)

According to Findeli (2004, as cited in Godin and Zahedi, 2014), RfD aims at helping, guiding, and developing design practice. It is a process of discovering and exploring for creation and production. Therefore, it is a type of research commonly adopted by professional designers, undergraduate design students, and artists. Glanville (1997, as cited in Megahed, 2017) describes RfD as research where the researcher is outside, looking outward.

The researcher conducting RiD makes investigations to produce knowledge and contribute to academia and literature. Frayling (1994) describes it as the most straightforward and the most common type of research in the design discipline that investigates social, economic, political, ethical, cultural, technical, material, and structural issues from historical, theoretical, and aesthetic perspectives. This type of research often draws upon established methodologies from the humanities, social sciences, or natural sciences (Fraser, 2013, as cited in Megahed, 2017). It is widely adopted by academics, master's, and Ph.D. students in universities and research centers to document objects, phenomena, and the history of design (Findeli 2004, as cited in Godin and Zahedi, 2014). Glanville (1997, as cited in Megahed, 2017) describes it as research where the researcher is outside looking inward.

3.2 My Approach

Day after day, AI is increasingly penetrating and influencing different aspects of our daily life and we encounter diverse variations of its applications. Designing a digital product consisting of that technology, especially in the health context, is a challenging process regarding engineering, design, ethics, regulations, clinical validations, and even the project's financial maintenance. As designers or innovators face these challenges, the complexity of these processes can be staggering. Therefore, the main purpose of this thesis is to document these challenges within the scope of design and to shed light on the decisions I made, and the approaches I adopted as a designer and a researcher during the development process of the AI-powered oral health mobile application prototype.

Before choosing this subject, I had a different idea. After I noticed the gap in the literature and that there was no study evaluating oral health apps on the market from the design perspective, I thought of researching to explore their user experience and usability by involving users as opposed to existing studies. This approach would have also been congruent with my background and the way of researching I was accustomed to. I would have been able to externalize myself from the products by acting as an observer and to a certain extent, it would have worked. However, at that time I had been running an innovation project for the last two years. I had been amid a TÜBİTAK project and already conducted the first two design research to improve the design of the prototype. Choosing the first subject would have caused a missed opportunity to share a very specific experience I was having as an innovator, an entrepreneur, a health scientist, and a design researcher. After I decided to choose the design process of the app as my thesis subject, I was not sure how to approach it and which parts of it were relevant and valuable for a thesis in design. It took a long time to come to a conclusion. Looking back, I saw that what I did throughout the process from the time I formed the basis of the idea to the present was a representation of my multidisciplinary education and experience. The design had an undeniable impact on the hybridity of the process, and it needed to be told in its complex nature.

In the traditional approach, a scientist commences researching with a question, aims to explain a phenomenon, and explores its causes by using a systematic inquiry without any influence of value judgments. As Archer (1995, p. 6) points out, he/she has a "*scientific ideal*" to produce valid explanations that can be reproduced by others to gain "*the powers of prediction*". The conditions and affecting parameters must be constant for research to be reproducible and inductive. I first tried to evaluate the design research I conducted during the R&D process with a traditional approach due to my positivist reflexes and draft my thesis accordingly. I encountered three major problems: First, I was not exempt from the influence of the value judgments. Positioning myself as an independent observer would be wrong because I was an actor and the protagonist of this story moreover, the decision maker. Second, the conditions were not constant, on the contrary, they were constantly changing. When I faced problems, there were times when I had to make quick decisions due to the limited time and project budget. Although I drew on the systematic methods while designing the research during the prototyping phase, not every inquiry was predetermined, they were rather "*quick and dirty*". During the research, I sometimes acted as a facilitator, asking additional questions, and making comments, in fact, I became a participant. Lastly, although I cared about the validity of the knowledge I generated, the process itself was not reproducible. Other researchers might have benefited from the methods I adopted and used them in their own projects, but we could not reach a generalizable conclusion. As Archer (1995) states, science is concerned with an explanation from which it can be generalized by establishing a cause-effect relationship. My aim was neither to produce a scientific explanation nor to predict the outcome anyway but to improve the product's design.

After these inner discussions, I went back and looked at what I did in the process. I conducted three user research to enable non-experts to use the expert system. The main goal of each study was to make the product more usable. This type of research that designers routinely did was not so meaningful for the design literature per se, on the contrary, what made them meaningful was the process itself. Therefore, these three studies should have been told without taking them out of context and their meaning in the course of the project. Moreover, in Zimmerman's words (Zimmerman et al, 2010, p. 310), my situation was "*messy*" due to everchanging conditions and there was "*a conflict of agenda*" because I designed the subject of these studies. The most appropriate approach in which I could position myself as an actor and explain the process from my perspective was *Research through Design*.

When I was formulating the idea, I was not deliberately and systematically "*conducting*" research but researching and peering around while in the prototyping process, I then conducted two design research with only a goal in my mind to improve the product's usability within the scope of the interaction between the expert system and the non-expert user[6], hence its performance. Therefore, I will discuss the ideation phase and three systematically conducted research in the next chapters by adopting the RfD approach. While this approach represents the type of research I have done during the process, it also acts as the material for the RtD approach I have adopted in the thesis.

---

[6] To be more specific, non-experts in dentistry.

In the final research, I had the prototype tested with the same goal by participants from three different backgrounds, which enriched the data. However, in the third study, there was another question that I wondered as a researcher: Which elements did the participants from different fields of expertise refer to most when evaluating the expert system? Where do their views intersect and where do they not? The answer to these questions obviously could not be addressed within the scope of *Research for Design*. They were theoretical questions and fall into the *Research into Design*. While thinking about this, all the pieces fell into place. I decided to adopt a hybrid design research approach, getting inspired by Frayling's (1994) classification of design research, and created a schema (Figure 3.1) by positioning the types of research he describes according to the actions that fit in my process.

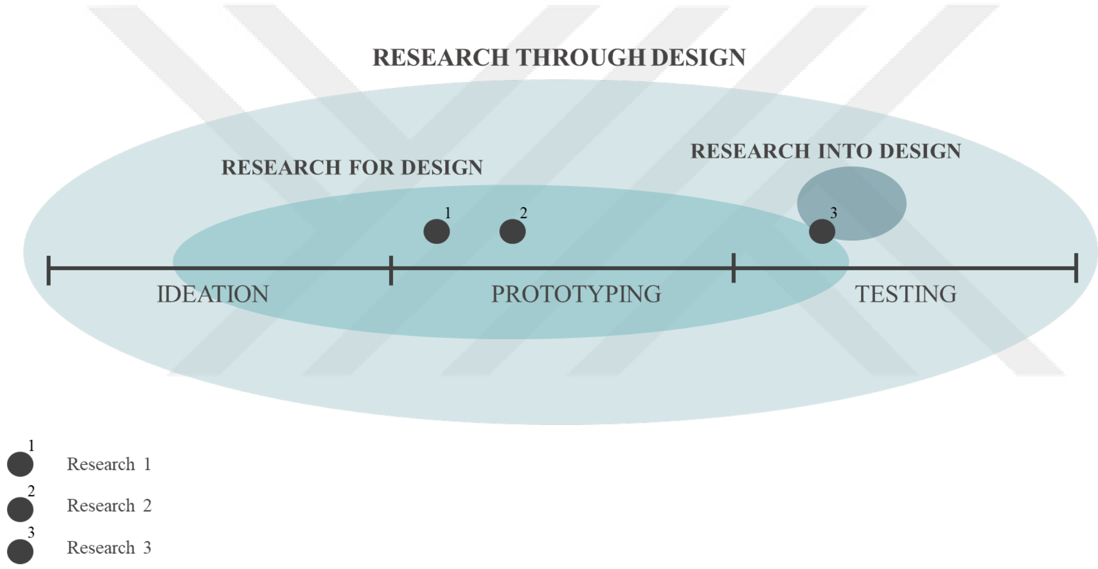


Figure 3.1: Types of design research in my process according to Frayling's classification (1994).

It would be necessary to remind some readers that I adopted RtD not to investigate the cognitive design activity, where the designer acts as a means. In the following sections, there will be some parts in which I leverage existing design paradigms to enlighten several issues in the ideation phase, but my aim is never to address this research field. Because it requires an extensive investigation in which the designer's thoughts are rigorously recorded and examined. RtD is rather the approach I adopted to document my design process in which first and foremost the design, then mDentistry, AI, human-computer interaction, and entrepreneurship took part.

Due to my position as a researcher constantly changing in the process, I adopt a hybrid research approach in this thesis. This hybrid approach not only encompasses all of Frayling's research classes but is also strongly influenced by the principles of action research.

4. MY DESIGN PROCESS

4.1 Ideation Phase

In 2019, I was introduced to an AI-powered app called ADA Health®. The app enables users to enter their symptoms and get information about possible diagnoses. I was impressed. I even wanted to work for them, so I checked their job postings. Since I was not a medical doctor, I tried to find a way to fit in. I have been studying design for several years, but I had no experience. During my Ph.D., my field of interest was oral medicine, which is a branch of dentistry focused on the manifestations of systemic diseases in the oral cavity. I was thinking of how I was going to answer this classic job interview question: "How will you contribute to our company?" I thought maybe I could help to develop an algorithm for oral lesions related to rare diseases. For a while, I worked on that but my application as an intern was, unfortunately (!) unanswered.

In 2020 March, I was about to move to Berlin. My equivalency process was towards the end and then unexpectedly the COVID-19 pandemic emerged, and borders were shut down, so I had to postpone my moving plans. During the quarantine, I received many messages from friends, friends of friends, and relatives about their oral health problems. They were telling me their complaints. I was asking questions and trying to form an opinion about the possible diagnoses. In the obscurity of the disease, all dental clinics were closed. Telling them to consult their dentists was not an option. I said to myself if ADA Health® was able to provide possible diagnoses of systemic diseases, and if I was able to form an opinion through messages, there should be an artificial intelligence that mimic my way of thinking. After all evidence-based dentistry exists. It should work out.

I did not know how to start. I was neither a software developer nor an AI specialist, but I had entrepreneurship experience. Back then in 2016, I started my design journey by taking a course called "Design Thinking" as a guest student at the university where I am now writing my thesis. The course was given by Prof. Dr. Hümanur Bağlı and we, the students taking the course, were expected to use design methods to improve the projects of the start-ups that we were randomly matched with. Luckily, the start-up I was going to work with was designing a medical device. I conducted interviews and questionnaires with stakeholders such as doctors, dentists, a pharmacist, patients, and their relatives to investigate factors that affect the feeling of trust towards a medical device. The device was a point-of-care testing device in the prototyping phase aiming at enabling patients to measure their coagulation time at home. The project owner, Erhan Ermek, was impressed with the results, so he offered to work together. In the following semester, I even presented my final paper at the 4T Design and Design Society Symposium in 2017. Later, Erhan applied to "*Grants4Apps*" organized by Bayer, and we were the only start-up from Türkiye that was invited to the investors' meeting held in Berlin. One of my responsibilities was designing the presentations for all the applications we did to incubation centers and other events. I worked with him for two years and it was a life-changing experience.

Based on what I learned from this experience, I knew the expectations of incubation centers from early-stage start-ups, so I started to build my idea by preparing a presentation where I explain how I define the problem, what I offer as the solution, how it will work, who the stakeholders are, what the competitors do, etc. With the preparation of the presentation, the idea took shape, and I wanted the name of the initiative to represent a motto that aims to facilitate users' access to scientific oral health information with the developing technology of artificial intelligence and empower them. Therefore, I named my start-up "my.mouth®".

I applied to the two incubation centers with a rough idea outline. The first one was ITU Çekirdek. I was evaluated in 5 minutes and was rejected. Acıbadem Incubation Center (AIC), which only focuses on the initiatives on the axis of health, spared an hour to understand the idea and my purpose. In September 2020, I was accepted to the "*BiggHealth Program*" by AIC which aims to train initiatives in the ideation phase to apply to TÜBİTAK 1512 Techno-initiative Capital Support Program. As part of the training program, we participated in various presentations about how to write the project and make a presentation for TÜBİTAK, legal issues such as regulations and IP rights, finance and business models, and one-on-one mentoring meetings. In this process, the idea was found too far-reaching for an early-stage venture, and its scope needed to be scaled down. For this reason, I decided to turn it into a mobile application in the field of pediatric oral health for use by parents. The training process proceeded through the preparation of the necessary form to apply for the grant program and each stage was controlled by the staff at the incubation center. 25 startups that were accepted to the training program at that time presented their project ideas to Acıbadem's board of directors and after the elimination, only 8 startups, including “my.mouth®”, were eligible to apply for the grant program. In 2021 June, my project was deemed worthy of a grant by TÜBİTAK, and the venture was established as a limited company in Pamukkale Technopark.

Before going into details about how I formulated the features, it would be pertinent to mention their intended functions.

4.1.1 The features of the app

There are seven features planned to take place in the app:

1- Symptom Assessment System (SAS)

2- Oral Examination Guide (OEG)

3- Emergency Guide (EG)

4- Dental Recording System (DRS)

5- Oral Hygiene Tracking System (OHTS)

6- Oral Health Products Platform (OHPP)

7- Nearby Dentists (ND)

The Symptom Assessment System is an expert system working based on scientific literature that assesses the child’s complaints, the child’s, and parents’ general health data such as systemic diseases, medications, dietary habits, etc., and provides users with the possible diagnosis of the complaints and future risks. The feature’s main purpose is to inform parents with tailored information about oral diseases, their etiology, the clinical signs, protective measures, and treatment, and the scientific references of the given information.

The Oral Examination Guide aims to teach parents what is normal and healthy in a child's mouth to distinguish the abnormal, the basics of dentistry such as the difference between primary and permanent teeth, when they erupt and shed, and how the users can make an oral examination properly.

The Emergency Guide aims to guide caregivers in emergencies such as trauma to minimize the negative impacts and refer them to a dentist as soon as possible. Parents' and teachers' awareness of oral traumas plays a critical role in handling the emergency until the dentist intervenes. The feature also cooperates with the Symptom Assessment System which can provide possible diagnoses of traumatic conditions.

The Dental Recording System aims to digitalize all dental records and make them easily accessible. It is also important to use these records for the Symptom Assessment System to make a more accurate diagnosis.

The Oral Hygiene Tracking System intends to assist parents in monitoring the oral hygiene habits of their child and introduces them to age-specific oral hygiene methods to improve the child's oral hygiene practices, keep records, and prevent erroneous habits. In the future, the app will be able to connect to the child's smartwatch, so that the system will remind the child when to brush their teeth.

The platform aims to inform parents of pediatric oral health products on the market and to raise awareness about the importance of fluoride use.

The app aims to inform parents of possible diagnoses and future risks and then directs them to a nearby dentist. They can inquire about dentists' specialties, their clinics, treatment prices, and patient reviews with the help of the Nearby Dentist feature. It encourages parents to take action and encourages preventive dentistry.

4.1.2 How I formulated the features

As mentioned earlier, ADA Health® was one of the factors that inspired me to come up with the Symptom Assessment System. It encouraged me that it is possible to bring artificial intelligence in healthcare into public use through mobile phone technology. However, there was no example in oral healthcare.

As a specialist dentist, throughout my career, I have seen a lot of patients come to the clinic by learning false information from the internet or their friends. Dentists do not enjoy dealing with these patients. Because some of them can be really difficult to convince. One of my colleagues used to call them "*smart aleck patients*", "*they are worse than illiterates*". Of course, I also had a hard time with patients who believed that fluoride caused pituitary calcification, but I have always believed that there were two sides to the same coin. Some people tend to believe in contrarian ideas, especially in healthcare, such as those about how influential big pharmaceutical companies called "Big Pharma" are in healthcare and how they are taking control of most doctors. The same accusations are being made against toothpaste companies. The ones who spread fake information are responsible for sure. However, I think we, the healthcare professionals, should ask ourselves: "Why do these people not trust us?", "Why is there such a thing as alternative medicine?" and "Why do some people choose to brush their teeth with baking soda and not toothpaste?". There may be many reasons, but I think two of them are very effective in these cases. First, to a significant extent, societies have limited or no access to scientific resources and it causes a huge gap between the public and science. Throughout the pandemic, we have seen people burning down vaccine centers, but have we seen any TV shows where pro-vaccine scientists have discussions with anti-vaxxers? Not that I know, at least in Türkiye. Second, there is sometimes a discrepancy between the diagnoses made by two different dentists/doctors. At the time when I worked as an oral surgeon, I saw frustrated patients who were told by the previous dentist that they would lose their teeth in five years, just because he/she saw a 7mm gingival recession. I told these patients it was not the case; they would not lose their teeth and there was a treatment for this type of recession. They felt relieved but confused at the same time. How come two dentists could tell them completely opposite things in a week? Which story should they believe in? A horror or a happy ending story? This is one of the reasons why some people experience distrust towards doctors. I believe this discrepancy harms our profession's respectability and I wanted to do something about that.

Diagnosing an oral disease isn't always easy, but most of the time it's not that complicated because, like any disease, oral diseases also have diagnostic protocols that show practitioners which steps they should follow during diagnosis. After graduation, very few dentists make an effort to update their knowledge. On the other hand, doctors know very little about oral diseases even though nearly half of the world's population suffers from them. The reasons I listed here represent how I define some aspects of the problem.

As a solution, I believe that an AI system constantly updated by the scientific literature can help dentists in their diagnosis and treatment decisions, assist doctors in detecting oral diseases in places where there is no dentist, collect data for researchers and policymakers, and improve the public's oral health literacy by providing information that can be easily accessed and understood. This is the goal that I am willing to achieve ultimately (Figure 4.1).

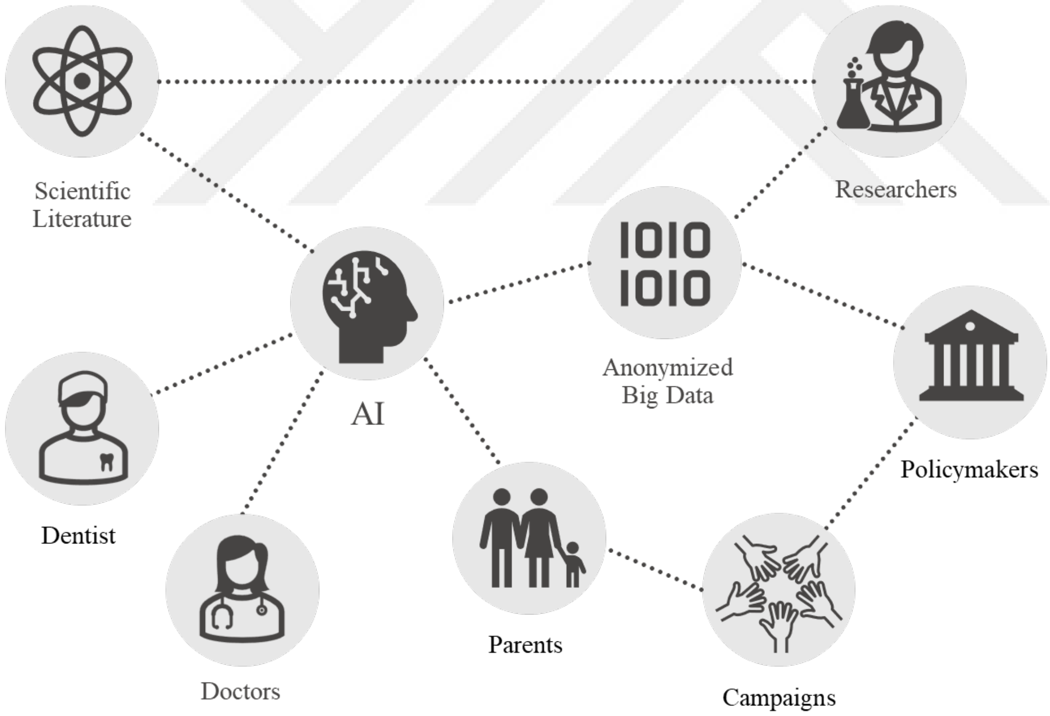


Figure 4.1: The ultimate goal of the project

As mentioned earlier, expert systems are generally designed for use by experts to support them in the decision-making process. In medical literature, these types of systems are called clinical decision support software. However, the system I developed targets parents, i.e., non-experts. At this point, the Oral Examination Guide aims to bridge the gap between the expert and the non-expert by providing not only how to examine a child's mouth but also the basics of dentistry, oral anatomy, and so on. There is no other example of such a feature in other apps. However, the idea was based on an obvious question: How will the app be used by the user without knowing anything about dentistry? So, I thought users needed more than a manual.

It is unthinkable to develop an app for children's oral health neglecting traumatic events. Since the development of motor functions is not completed in early childhood and sports activities are performed frequently in the later periods, traumatic lesions prevalently occur. Therefore, it is vital to know how to handle the situation prior to consulting a dentist. There are several apps for this purpose on the market. For example, the ToothSOS® application provides users with real-case photos of injuries that may occur as a result of trauma and makes recommendations for both professional and general users. With the same purpose, I wanted to add the Emergency Guide working in cooperation with the Symptom Assessment System.

One of the most common problems encountered during the examination is the inability to access data and records regarding the patient's history. Patients often have difficulty remembering the previous diagnosis and treatment. However, these data are important in preventing misdiagnosis and future complications (WHO, 2016). Türkiye is one of the countries that has made significant progress in the digitalization of health data. The citizens can access their personal health information (the data about diagnosis, treatment, test results, etc.) through the E-Nabız® app. However, it only keeps data about the health services provided by social security. According to The Health Statistics Yearbook 2018 of Türkiye, the annual oral health expenditure in the country in 2018 was around \$ 2 billion, and half[7] was spent in the private sector, which tells us the same amount of oral health data is held in private clinics. It is not digital and portable. The Dental Recording System appeared to solve this problem.

One of the major reasons for oral diseases is poor oral hygiene. Numerous mobile applications on the market aim to assist users in brushing their teeth to enable them to brush every tooth surface in the oral cavity for a sufficient time. The most widely used of these apps are those that connect the electric brushes produced by the big

---

[7] For the sake of clarity, the expenditure stated in Turkish Lira has been converted to US dollars, taking into account the US dollar-Turkish lira exchange rate of the period.

toothbrush manufacturers via Bluetooth. Disney® characters and gamification elements were even used in the application designed by Oral-B® for children. However, very few of the existing applications have educational content in which brushing techniques and tips according to age are explained to parents. The Oral Hygiene Tracking System aims to keep track of oral hygiene habits and remind parents and children to brush while providing educational content and giving tips.

Which product to use is a frequently asked question to dentists. I thought it would be helpful to show existing oral hygiene products for children on the market to parents. Besides toothbrushes and toothpaste, other products such as interproximal brushes or tooth mousse are known by very few. The Oral Health Products Platform aims to provide information about their ingredients, functions, and way of use. Another intention in introducing this feature was that I thought maybe it could be used as a sales channel.

The last feature, Nearby Dentist, was not even in the project description in the beginning. Throughout the mentor meetings in the BiggHealth Program every time I told the functions of the Symptom Assessment System, I was asked, "so, what?", "As a user, what am I going to do next?" Frankly, I was first annoyed by these questions. Was providing tailored information not working enough? However, the idea that the user may want to be accompanied on their entire journey started to make sense after a while and I decided to add this feature.

4.1.3 How I planned the TÜBİTAK project: From an idea to a prototype

As mentioned earlier, the training program held in AIC was mainly focused on filling out the "AGY112-01" form which consists of 10 main sections with subsections: information about the entrepreneur, a brief overview of the idea, information on technological verification to be performed under the project plan, personnel information to work on the project, information about the type of the company to be established, the market analysis, marketing strategy, production plan, financial plan, estimated project cost forms, and personnel expenses estimated cost forms. The duration for preparation of the submission took nearly four months and the application file has been revised six times.

In the form, it is expected the entrepreneur elaborate on their idea from a technological, business, and financial perspective. In addition to the product overview, the aim of the project, and its societal impact, I had to specify the aspects of novelty, what kind of technology I would leverage to build the prototype, the scientific literature on which the project was based, which methods I would adopt, tangible outputs to be achieved, and success criteria for these outputs. I also had to define each process in the project under a work package where the steps were explained in detail and shown on a work-time graph. For the business part, I prepared a very detailed file where I defined the targeted customer, rivals, and the business model; set the price for each targeted market; made a five-year plan where I estimated the targeted number of customers for each year; determined the channels through which I would reach that number; set marketing, and advertising strategies and estimated their expenses; estimated the future income and the size of the market share that the company would achieve with this income. Detailed tables on the expenses to be made for the project including equipment to be purchased, operating expenses, rents, taxes, and bills to be paid, as well as on which sources these expenses would be met from other than the grant and when the company would start to compensate for these expenses took place in the form.

I planned the project to last 13 months under three work packages. I allocated 4 months for the first work package which I aimed to develop the SAS algorithm and prepare the medical content for the features, 6 months for UI/UX design and coding processes, and the last 3 months for testing the algorithm (Figure 4.2). The work that I undertook to complete in the project: Preparation of medical contents for 4 features; development of the algorithm; designing mid-fidelity interfaces for all features; encoding of SAS in Android system; collecting real patient data by conducting clinical research; testing the performance of the algorithm with real case data and conducting user research. The success criteria were to complete at least 70% of the medical content, build a prototype to be tested, and achieve at least 70% successful algorithm performance.

I formed a five-person team with three staff members, including myself, and two consultants. I included the first consultant to assist in the development of the algorithm, the design, and the coding processes, and the second consultant to assist

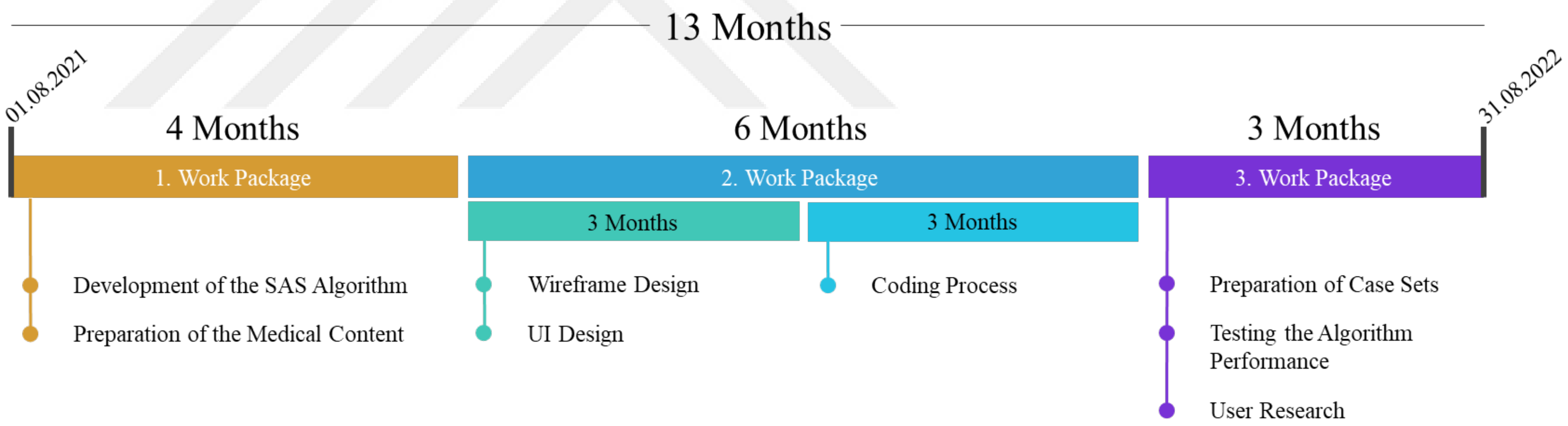


Figure 4.2: The estimated timeline of the project in the submission

with the design research. Because the project budget was limited, I decided to do all the work myself, except for coding and collecting patient data in clinical research.

I submitted the project documents in February 2021 to TÜBİTAK for a grant of approximately 200,000 Turkish Liras ($ 28,000), including personnel, consultant, and equipment expenses. On 26 June 2021, the results were announced, and I was asked to establish a limited company until the 15th of July. I applied to Pamukkale Technopark for an address required for incorporation procedures. The company establishment was approved after the board of directors meeting, so "my.mouth®" was incorporated on 8th July 2021.

4.2 Prototyping Phase

The project officially started on the 1st of August 2021. It took ten months to build a working prototype of SAS, the only feature I will focus on in the thesis.

4.2.1 The first work package: Developing the SAS algorithm

The SAS algorithm is simply an expert system that consists of the same components mentioned in the previous chapter. Unfortunately, I will briefly tell how I developed its knowledge base and inference engine due to confidentiality, therefore as expected its explanation facility and user interfaces will be at the center of the discussion.

During the first three months, I created a medical library by conducting an extensive literature review, in which oral and dental diseases in children and the systemic, genetic, and environmental parameters associated with these diseases were added. I divided the diseases cluster in the library into oral and dental diseases and other parameters affecting these diseases and classified oral and dental diseases cluster (120 diseases/conditions) as 8 main titles and 10 sub-titles according to their etiology and clinical features. I scanned an average of 500 scientific articles using search engines such as PubMed / Research Gate / Science Direct and read 386 of these articles and found 214 suitable for creating the knowledge base. I extensively benefited from a book on pediatric dentistry (Koch, Poulsen, Espelid, & Haubek, 2017) and collected related data for each disease in the medical library. While collecting the data, I paid attention to the scientific strength of the articles on diseases and the published journals and gave priority to meta-analyses and systematic reviews on relevant topics. When I could not find articles with high scientific power about the related disease, I marked the data obtained. Finding the right source was harder than extracting the data.

I examined the clinical signs of 120 oral and dental diseases in the library and decided that 4 of the diseases were not in the field of pedodontics, but rather the field of pediatrics. I determined that data that cannot be detected by the user is needed for the diagnosis of the other 6 diseases. For these reasons, I added the first 4 diseases to the list of etiological parameters; and deactivated the other 6 diseases. Considering the collected data, I classified etiological and risk factors (308 in total) of 110 oral and dental diseases under 17 main headings. I grouped 107 symptoms of oral and dental diseases in the library under 9 titles as "*Symptom System Version 1*" (SSV1). I prepared questions concerning the symptoms of SSV1 and possible answers for multiple-choice questions.

Since the algorithm will start by asking the user what his/her complaint is, I designed low-fidelity wireframes considering SSV1 by using the Figma program. During this process, I sought the answer to the question of how users find the right symptoms and I ascertained that the SSV1 was problematic in terms of user experience. The symptom system needed an extra algorithm called Natural Language Processing (NLP) that could understand the users' descriptions of symptoms and associate them with the symptoms in the library by asking "*Did you mean that?*", to which we are all accustomed in Google Search. However, developing the NLP algorithm was a task that should be done within the scope of a separate project. Owing to overcome this shortcoming, I decided to conduct user research to investigate how non-expert users describe oral diseases. I built the SSV1 based on the symptom descriptions in the scientific literature. I thought that if I could identify how non-expert users describe these symptoms, I could find a way to facilitate them in selecting relevant symptoms and bring the expert system built in medical terms closer to non-expert users.

4.2.1.1 Research 1: Understanding how non-expert users describe oral diseases

After identifying problems in SSV1, I decided to conduct user research to collect data about how non-experts describe the symptoms of oral diseases. I thought I could draw upon their descriptions instead of medical terms to make the symptoms easier for users to understand.

I collected photos of 40 different oral diseases and conditions included in the algorithm and photos of healthy cases taken from the same angle via Google Images to enable participants to compare. I paid attention to selecting less disturbing case images and prepared a presentation using the PowerPoint program, including information (number of deciduous teeth, number of permanent teeth, their position in the mouth, where the midline is located, etc.) that was planned to take place in the OEG. Five participants (Table 4.1): two parents, two teachers, and a designer agreed to participate in the study. The two most important reasons I included these participants were that they were all non-experts and were people who could devote their time to my study. During the interviews via Zoom, I asked them to describe these images as if they were talking to a dentist over the phone. In the meantime, I took audio and video recordings. All interviews were conducted in Turkish[8].

While 3 participants (two teachers and a designer) participated in the interviews individually, two participants who were parents participated in the study together which allowed them to discuss the cases. It was interesting to witness the influence of their background in the way of expressing their descriptions. For example, due to the reflex of empathizing with the user in the design approach, the designer made his descriptions as if he were a parent and talking to a dentist. On the other hand, one of the teachers (P1) occasionally said, “The parents would probably describe it this way.” Since the parents were experienced in the dental development of their children, the effect of being a parent was apparent in the inferences they made.

During the interviews, I did not act like a typical observer. On the contrary, I got involved in the discussion where necessary and guided the participants. Because there

[8] Original versions of the transcripts are in the Appendix.

were some limitations in the design of the study due to its nature. One of these limitations was that the cases in the mouth photographs were not isolated. Therefore, when they could not find the relevant area, I had to direct the participants to the points I wanted them to describe, and at the end of each description, I asked if the definition I used in the algorithm could be suitable for them. The purpose of this research was to learn how participants would describe the anomaly they saw, rather than whether they could detect the anomaly in the photographs. When evaluating the results, in addition to what they said, I also paid attention when I needed intervention, which became another type of data for the study.

Table 4.1: Backgrounds of the participants

Participant No.	Gender	Age	Background
P1	Female	37	Music teacher
P2	Male	37	Design Academician
P3	Female	37	Former Turkish teacher (She has been unemployed over the past 5 years.)
P4	Female	51	Parent (Mother) – Chemical Engineer
P5	Male	56	Parent (Father) – Former small business owner

4.2.1.1.1 The results and changes in SSV1

I can list the results and the changes in SSV1 I planned to make as follows:

- In general, the participants described the cases in terms of color, shape, size, and surface properties which caused me to pay attention to these parameters when updating the symptom definitions.

- I had defined "*caries*" by the parameters of "*substance loss*" and "*discoloration*" in SSV1. I decided to use "*caries*" itself as a symptom in the "*Symptom System Version 2*" (SSV2), the next version of SSV1, because the participants used the word "*caries*" in their descriptions.

- Since the "*substance loss*" parameter in SSV1 was not adopted by users, I changed it to a title rather than a symptom and converted it to "*substance loss in dental tissue*" with the contribution of P2. In the subsection, I added "*caries*", "*tooth wear*", "*fracture*" and "*cracks*" as symptoms.

- Since the location parameter in descriptions is frequently used by the participants such as "*tooth discoloration*", "*gingival redness*", and "*jaw crookedness*", some symptom clusters were classified according to locations such as "*the tooth-related symptoms*", "*the soft tissue- related symptoms*" and "*the symptoms related to the alignment of the teeth and jaw relationships*".

- There is a physiological difference between enlargement and swelling. While cells proliferate in enlargement, swelling is caused by a buildup of fluid in the tissues. For example, swelling in the tissue is seen in the inflammation of the gums (gingivitis), while gingival enlargement can be seen in patients using some anti-epileptic drugs. However, it would be unfair to expect non-experts to discern this difference. Since the participants used the words "*swelling*" and "*enlargement*" in the same sentence in the definitions, in SSV2, I decided to combine the symptoms planned separately in SSV1.

- P1 described the cases (Figure 4.3) as "(a) The surfaces of the teeth are jagged like a canal or a stream…, (b) the tip of the tooth is indented like a crater." P2 described the image (b) as "This tooth looks just like how plastic looks when it melts. The tooth is out of the geometry it should have." P3 said about the image (a), "There is deformation on the surface of all teeth… I would say congenital enamel damage." P4 stated, "The teeth have the appearance of a crater…I think the most appropriate definition is definitely integrity deformity and even surface roughness." Based on these descriptions, I changed the symptom of "dental tissue disorder", which I set to describe defects in conditions such as amelogenesis imperfecta, or enamel hypoplasia, to "*Rough Indented Surface in Enamel*". I decided to place the symptom under the "*Disturbance in tooth tissue and surface*" title.

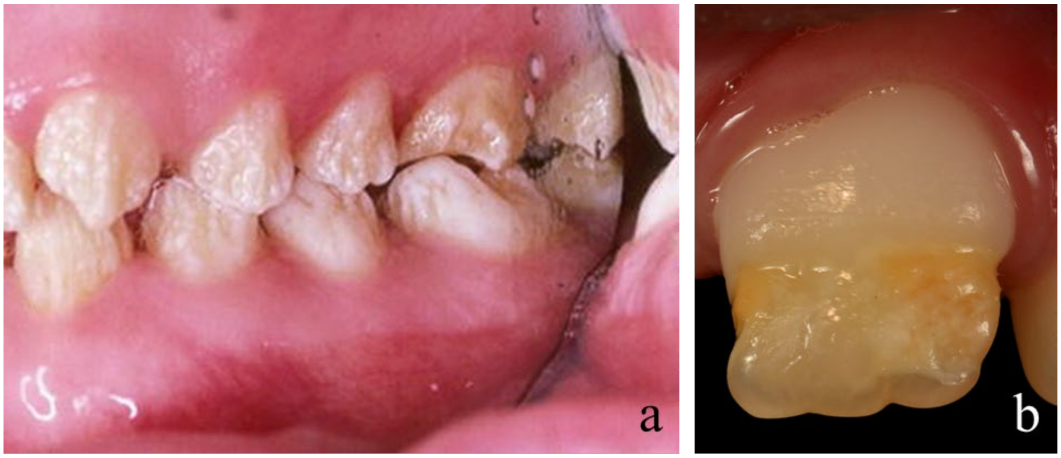


Figure 4.3: Images of amelogenesis imperfecta[9] (a) on the left and trauma-related enamel hypoplasia[10] (b) on the right.

- Orthodontic problems are usually defined by the aberration in the contact points of the teeth or the lack of contact. With this reflex, I first tried to describe the symptoms related to the alignment of the teeth in SSV1 through the contact of the teeth. For example, the case in Figure 4.4 is normally defined as "*the incisors that do not come into contact with the occlusion.*", in short, "*open bite*". However, I predicted that these definitions would be too complicated for the user. During the interviews, after letting the participants describe the case, I asked them if "*open bite*" makes sense.

> P1: There is a gap between the teeth like Madonna, only the place of it is different. (Does the open bite mean anything?) No. Open bite does not make any sense.
>
> P2: The child's jaw is not fully pressed. Front teeth stay in the air. (Does the open bite mean anything?) No. It does not.
>
> P3: There is a very symmetrical gap between the teeth. An interdental space… The space is symmetrically between the upper and lower teeth. (Does the open bite mean anything?) No, it does not make any sense.

[9] Copyrights Unknown. For the source, look at Url-1.

[10] Copyrights Unknown. For the source, look at Url-2.

P4: The teeth are not fully closed. There is even a space. (Does the open bite mean anything?) No, I would not get it.

P5: The teeth are not fully closed, and the open bite does not say anything to me. It's like it's flattened from the sides and protruding from the front.

Figure 4.4: An image of a case of open bite[11].

By considering these definitions, I determined that users see space rather than contact, and I prepared a separate section for symptoms related to space under the "*the symptoms related to the alignment of the teeth and jaw relationships*" cluster. Thus, I created 3 separate symptom titles according to where the gap is: (1) "*Space between two adjacent teeth*", (2) "*Spaces between the teeth in the upper and lower jaws in the anterior region when the jaws are closed*", and (3) "*Spaces between the teeth in the upper and lower jaws in the posterior region when the jaws are closed*".

- Another orthodontic problem that I had difficulty with translating its definition for non-experts was "*deep bite*" (Figure 4.5) which is technically described as "*the increase of overlap of maxillary incisors over mandibular incisors.*" After letting the participants describe the case, I asked them if "*deep bite*" makes sense.

[11] Copyrights Unknown. For the source, look at Url-3.

P1: The direction of the teeth is wrong. The teeth are tilted back if there is no deficiency or difference in size. (Does the deep bite mean anything?) In this case maybe. Because the teeth had covered the others here.

P2: The middle teeth of the lower jaw have not erupted. (I stated that there are teeth) When the jaw is closed, the middle teeth of the lower jaw are not visible. (Can it be said that the upper teeth have covered the lower teeth?) It can be said yes. (Does the deep bite mean anything?) No. I think we need something more descriptive. Longer definitions are needed, not just one or two words.

Figure 4.5: An image of a case of deep bite[12].

P3: Upper teeth overlapping lower teeth, lower teeth not visible. (Would deep bite mean anything?) Yes, a bit when you asked by showing this.

P4: There are no lower teeth, or they were closed on it? (I say there are teeth) When it is closed, the lower teeth are not visible. Oh, I would check that the lower teeth are not visible, I would select the relevant symptom, but the deep bite does not mean anything.

P5: Impairment of the occlusion, the teeth overlapped. (Does deep bite mean anything?) No, I wouldn't say, I would say there is a malocclusion.

Based on these descriptions, I changed the "*deep bite*" symptom to "*Upper front teeth covering/overlapping lower front teeth (Deep bite)*".

[12] Copyrights Unknown. For the source, look at Url-4.

- There was a misconception about gingival recession among the participants. When I showed them a case of the gingival recession of a single tooth (Figure 4.6), they stated as follows:

P1: There is a recession in the gum. The tooth seems to fall out at any moment. It's like the gingiva is elongated towards the roots. The flesh looks at the bottom of the tooth like gum denudation. (When I ask if she would select the gingival recession) When you say gingival recession, it's like a slight recession in all teeth, there is a weirder situation here. Too overexposed. As if there was a millimetric measurement here, I would say it more comfortably.

Figure 4.6: An image of a case of gingival recession[13].

P2: The root of the tooth is out and visible. (Is the gingival recession a suitable term for you?) I would pick the gingival recession when there is more than one recession. I would think that if there was a recession, all of them would recess. It's like the root is exposed.

P3: Exposure of the root. Gingival loss. (Would gingival recession be suitable?) Yes, I would select that.

[13] Copyrights Unknown. For the source, look at Url-5.

Based on these descriptions, I changed the "*gingival recession*" symptom to "*gingival recession & loss/ root exposure*".

- Geographic tongue results from the loss of tiny hairlike projections (papillae) on your tongue's surface (Figure 4.7). For ease of understanding, I defined the related symptom as "flattening on the tongue" in SSV1. However, P2 stated "I would say that red dots and white spiral patterns have formed on the tongue. (I mentioned the form of papillae and asked if there was a difference between those two fields) I would look at my own tongue. I don't know which one should be, but it shouldn't be both. (Would flattening on the tongue make sense?) No, It wouldn't. I would say that there are some patterns on the tongue. I guess I couldn't evaluate correctly because I don't know which one is correct. But I'd say it's a problem that it's so heterogeneous." P3 described it as "Trippy patterns on the tongue. White patterns on the tongue. Loss of projections in the tongue. Smoothing and flattening of the tip of the tongue. That roughness is gone." P4 said "Normally there should be a texture on the surface of the tongue, it looks like it had disappeared. I understood what you meant by flattening on the tongue, but I think the pattern makes more sense. The surface of the tongue normally has hairlike structures and they have disappeared." Because of the frequent use of the word "pattern", I changed the name of the symptom to "*Flattening and patterns on the tongue surface*".

Figure 4.7: An image of a case of geographic tongue[14].

[14] Copyrights Unknown. For the source, look at Url-6.

- Normally when the jaws are closed, the incisal and occlusal ends of the upper teeth contact the outer side of the lower teeth. The reverse of this situation is defined as "*crossbite"*. When it occurs in the anterior region, it is called "*anterior crossbite"* (Figure 4.8). When it occurs in the posterior region, it is called "*posterior crossbite"* (Figure 4.9). In the SSV1, I used "*crossbite"* as a symptom regardless of its location. The participants described Figures 4.8 and 4.9 as follows:

> P1: (For Figure 4.8) This time, the upper teeth are inwards, falling behind the lower teeth, but it should be the other way around… (For Figure 4.9) There seems to be a problem with the orientation of the teeth or the jaw. It's like the upper teeth are protruding behind the lower teeth. There is a gap between the upper and lower teeth. I don't know if it's about the jaw, its position, or the size of the teeth.

> P2: (For Figure 4.8) The upper jaw stays inside the lower jaw, is this an inward bite?... (For Figure 4.9) The teeth between the canines are in the air. Jaws do not press together. (What do you think of the back teeth?) Now when you say that, it seems like they're a little inside. Is it correct?

Figure 4.8: An image of a case of anterior crossbite[15].

[15] Copyrights Unknown. For the source, look at Url-7.

Figure 4.9: An image of a case of posterior crossbite[16].

P3: (For Figure 4.8) Lower jaw is forward. The lower jaw covers the upper teeth. (Would the crossbite mean anything?) Yes, it would. It might be more descriptive if the jaw term is added. I'd understand if you'd say the jaw is turned upside down… (For Figure 4.9) Short upper teeth, shortness of the upper four teeth. (Is there anything in the back?) The color and texture did not seem different because it was normal. Is the number different? (No, is there a problem with the jaw relationship?) No, I couldn't say anything.

P4: (For Figure 4.8) In occlusion, the upper jaw remains inside. Are the upper teeth visible when closed? They are not visible completely… (For Figure 4.9) It looks healthy to me (What about the teeth in the back?) It's like they overlapped. The upper teeth should have protruded a little more, as those in the healthy image.

P5: (For Figure 4.8) It's the opposite of what it was before, it's not a deep bite, but the opposite, I would say that the lower teeth cover the upper teeth as if the upper jaw is behind… (For Figure 4.9) The back teeth do not contact properly; the upper teeth seem to move inwards.

[16] Copyrights Unknown. For the source, look at Url-8.

Based on these descriptions, I changed the symptom "*crossbite*" to two different symptoms: "*Reverse closure of the upper front teeth relative to the lower front teeth (Anterior Crossbite)*", and "*Reverse closure of the upper back teeth relative to the lower back teeth (Posterior Crossbite)*",

- Since the images were useful for comparison, I decided to add relevant images and illustrations to the symptoms.

- It turns out that users need to know the anatomy of the teeth to understand that some conditions are abnormal. For example, in the abrasion case, P2 stated "Since I did not know the part under the gum, that is normal, I would not have thought to say that there is a loss of substance there."

I made some decisions in SSV2 independent of the study results. When examining patients in the clinic, they are first asked what their complaints are and then where they are. However, due to recurring symptoms such as bleeding or discoloration in soft tissue symptoms, I created a new cluster where the user first selected the location. So, I classified soft tissue symptoms under five locations: gums, cheek, lips, tongue, and skin.

While showing the images about orthodontic problems, I had to say from which angle the shot was taken. This showed that three-dimensional problems are not fully understood when they are represented as two-dimensional. Therefore, I decided that designing three-dimensional models in the future will improve the user experience and the performance of the algorithm.

A limitation of the study was that the photographs of the same case were not taken from different angles. The quality issue with some photos also affected the evaluations. Also, judging by looking at photos is not the same experience as judging by a mouth in reality. As P4 stated, "We are just looking at the picture, not the same experience, the child comes and says that she has pain or bleeding, we look at it and see. It's not the same as in the picture."

I can summarize the conclusion as follows:

1) I decided to remove or modify some symptoms in SSV1 given the new data. While I made some of the changes in the definition of the symptom, I merged some symptoms or added new symptom definitions. I called the revised version of the system "*Symptom System Version 2*" (SSV2).

2) While only the changes made in the definitions affected the explanation facility, adding or removing symptoms from the system necessitated rebuilding the algorithm. As a result, the user research conducted in this process provided information about the interfaces and the explanation facility and caused the inference engine to be redesigned.

3) Some changes, such as those I made in soft tissue-related symptoms, occurred independently of the user research results.

4.2.2 The second work package: Designing the user interfaces and the wireframes

After the first work package, I made the changes mentioned above in the algorithm and designed the SSV2 where I classified 129 symptoms under 9 main headings:

1) Tooth-related problems

2) Sensory symptoms

3) Soft tissue-related problems

4) Systemic symptoms

5) Functional difficulties

6) Trauma-related symptoms

7) Symptoms related to the alignment of the teeth and jaw relationships

8) Problems related to eruption and shedding

9) Saliva-related symptoms

Considering the working principle of the algorithm, I decided that the user flow should consist of four steps:

1) Creating the child's profile, where the user will enter the child's health data.

2) Selecting symptoms.

3) The assessment process, where the algorithm will ask extra questions to users.

4) The results, where the user will get information about the possible diagnoses based on the inputs and the child's health data.

The reason for the child profile was to prevent the algorithm from asking for the child's health data in each assessment and to shorten the process. I decided that the data obtained from the user in the profile was those that hardly changed and corresponded to the etiological risk factors.

In the second work package, I primarily focused on the design of the home page and the pages where the user will select the symptoms. Because I needed to test whether the symptom system I renewed could be used by users. I can summarize the design process of interfaces as follows:

- I designed the wireframes on paper and determined which information would take place where. The information architecture was already somewhat evident in relation to the algorithm. I just decided to what extent they should be included on the pages.

- Then, I designed low-fidelity user interfaces on Figma, so I could determine the components needed and their functions.

- I mapped the flow and the links between pages.

- While designing the mid-fidelity interfaces, I opted for using neutral colors and did not pay extra attention to aesthetics. There were two reasons for this: First, this was my first interface design, and I did not have the required skills. I learned the theory and research methods rather than how to design in my master's education. I took a UI course for a semester as a visiting student at Marmara University after submitting my project to TÜBİTAK, but it was not sufficient when compared to what a design student learned in their education. For this reason, I have never positioned myself as a designer. It was something I had to do during the project due to limited resources. Second, it was just a prototype. My priority was to design a working system that could be used by non-experts. A decent look was enough for me.

- Here are some of the decisions I made for the user interfaces of the symptom selection pages:

  a) For the profile page to be accessible on every page, I placed the icon with the child's photo in the upper left corner (Figure 4.10).

  b) I decided to design the symptom selection process in a way similar to the add-to-cart action in online shopping services that we are all used to. I have placed a medical chart icon that can be accessed on each page and allows users to see the selected symptoms before clicking the assess button. I also added a marker on the medical chart showing the number of symptoms added, just like a shopping cart (Figure 4.10).

  c) I've placed a section for the prevalent symptoms under the medical chart for easy access (Figure 4.10). I identified these symptoms by looking at the frequently seen symptoms in children who applied to the pedodontics clinic.

  d) Under each heading of 9 main symptom classes, I wrote some of the symptoms in these classes so that the user can guess what they may see there, even though they do not understand what the heading means (Figure 4.10).

  e) In UI design, inactive buttons are normally designed in pale color and user-selected buttons become bright. I, on the other hand, tried the opposite. The user would be able to see the selected symptoms in the medical chart, so I wanted to prioritize the ones they didn't select on the list, not the ones they chose. Therefore, inactive buttons would be dimmed when selected (Figure 4.11).

f) Another decision regarding buttons was to have other symptoms inactive to avoid selecting more than one symptom in mutually exclusive symptoms (Figure 4.11).

g) I designed illustrations of 20 symptoms related to the alignment of the teeth and the jaw relationships so that users can only focus on the anomaly. I also designed illustrations of normal occlusion to allow users to compare and placed an icon next to the image so that it can be viewed by pressing and holding the screen (Figure 4.11). As an example, I collected case images suitable for tooth-related symptoms and made them visible via the photo icon I placed next to the symptom buttons. However, there were a couple of things that worried me about case images. First, although it was not important for the prototype, the photos I will use in the application had to be copyrighted. In addition, I thought that the degree of cases shown in the photographs may mislead the user, whether mild or severe, rather than the situation in the child's mouth. Therefore, I decided to add multiple photos for each symptom that gradually show their intensity in the future.

h) Since the symptoms related to the alignment of the teeth and jaw relations are numerous, I decided to design the user flow differently for this class only. I gathered the symptoms in this class under 6 sub-headings and determined a flow where the user can see all the symptoms in 6 steps (Figure 4.11).

i) I thought the brightness of the forward button might act as a sign showing that the user must make a selection on some pages. I decided to keep the forward button inactive when no selection is made (Figure 4.11).

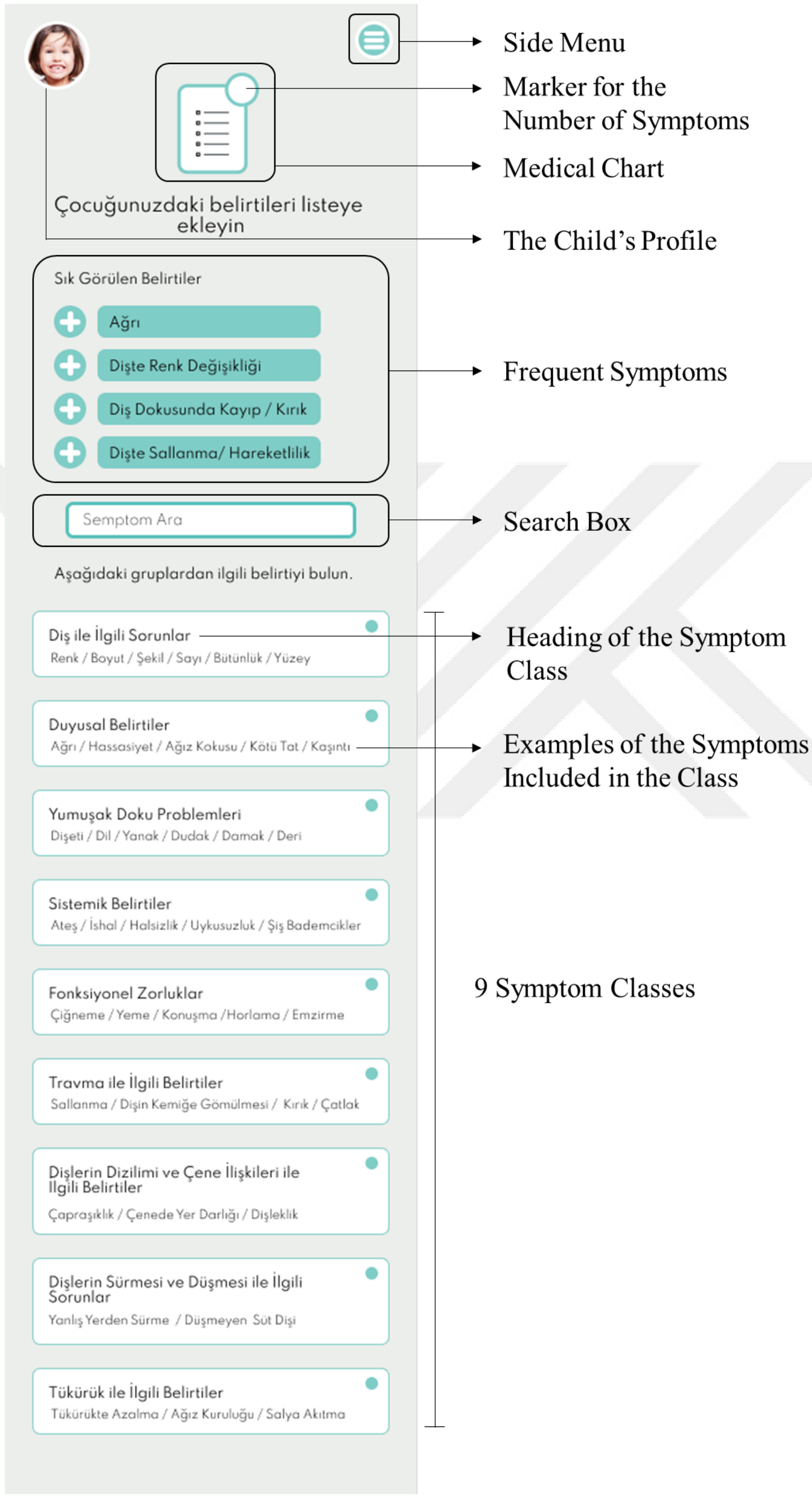


Figure 4.10: Symptom selection page.

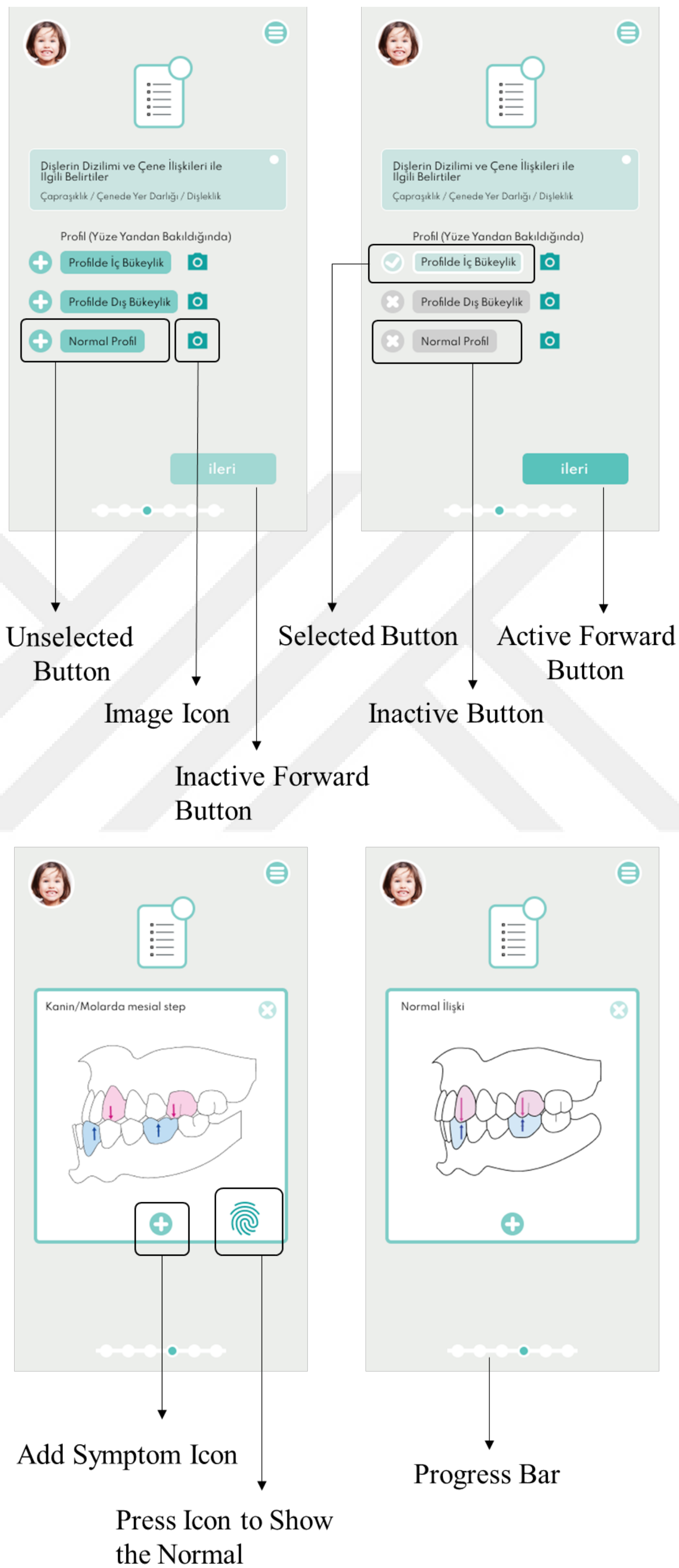


Figure 4.11: Symptom selection interfaces of the first prototype.

4.2.2.1 Research 2: Testing the two symptom selection prototypes

I rearranged the symptom list based on the data obtained from the first user survey. I designed the interfaces of the symptom selection pages in Figma. The prototype made with these interfaces had to be tested by the user. That way I would see how both the interfaces and the symptom list would be evaluated by users.

I designed two prototypes with two user flows in one of the symptom classes which is related to the alignment of the teeth and the jaw relationships, and I wanted to investigate how these flows in the prototypes would be evaluated by the users. In the first prototype, the symptoms were classified under 6 sub-headings as before mentioned where the user can see all the symptoms in 6 steps (Figure 4.11). In the second one, the total 23 symptoms in this section were sorted where the explanation and displayed illustration of each symptom took place on a single page with three answers: "*yes*", "*no*" and "*I do not know*" (Figure 4.12).

I included participants from the previous study (Table 4.1) and prepared 8 case stories including gingivitis, caries, geographic tongue, amelogenesis imperfecta, teething, a fracture, and two malocclusions for the study. In the stories, I determined the symptoms such as fever and bad breath that the participants could not understand from the photos to tell them. I asked them to identify other symptoms through photographs. Participants evaluated one of the two malocclusion cases using the second prototype.

I did not wait to complete the user research to make changes to the interfaces. After each interview, I applied the ideas that I found reasonable to the prototype. I conducted the interviews with the participants in the following order: P5 and P4 (together), P1, P2, and P3.

4.2.2.1.1 The results and the changes in the prototype

- Participants were able to find the majority of symptoms that were underscored in the case stories and that they identified in the images. Despite many shortcomings, this was enough to move on to the next stage in terms of the progress of the project.

- In the case where the tooth in the upper jaw was embedded in the bone as a result of trauma, some of the users could not find the relevant symptom. There were two reasons for this: first, the appearance of shortening of the part of the tooth in the mouth gave the impression that the tooth was broken. Here, the limitation created by an evaluation through photography may have led to this result.

- The second reason, the case was described as the appearance of the tooth embedded in the bone in the upper jaw, as the tooth rose and moved towards the nose. This showed the need for illustrations and even animations for trauma symptoms.

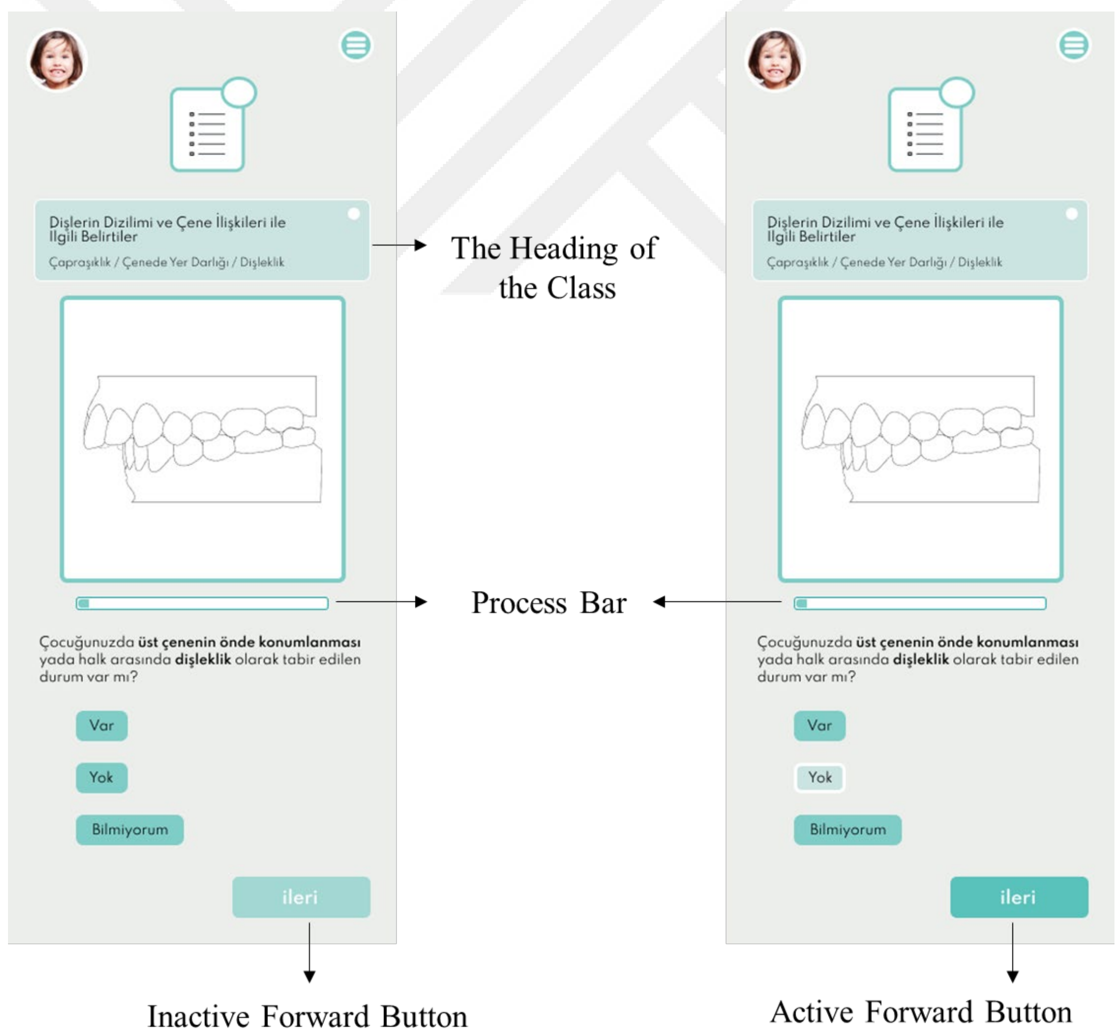


Figure 4.12: Examples of symptom selection interfaces of the second prototype.

- In some illustrations (Figure 4.11), users had to hold down on the icon to see the normal state. Since this action is done with the thumb for right-handed users, I preferred the fingerprint icon. However, this icon meant nothing to users. P5 said that there should be a phrase like "*show the normal*" on the icon. P2 on the other hand said that he didn't want to press the fingerprint icon as it was about security and gave the impression that the device would register his fingerprint. P3 stated that she was uncomfortable with the word normal by saying "It feels weird, so what my child is abnormal?" I changed the icon right away, but "*show the normal*" remained as I could not find another suitable word for it.

- I chose the camera icon as the icon for the button that shows the relevant image when pressed next to the symptoms (Figure 4.11). The reason for this was that the existing icons for the image included the mountain and sun symbols, giving the feeling of seeing the panorama photo when pressed. However, P2 also objected to this icon. He stated that it gave the impression that the device would connect to the camera on the phone if he pressed it. I worked on the image icon. Even though I designed icons with a tooth and a magnifying glass, I was not successful. For this reason, I decided to use the icon with the mountain and sun that everyone is used to.

- The headings of the classes (Figure 4.12) did not give the impression of going back to the main symptom page when pressed. None of the participants used it. P2 suggested using a component that acts as a directory to give this impression.

- P1 said that she entered the symptoms extra because she did not want to enter anything missing. Upon this situation, I asked her if she noticed that she did not need to select symptoms on some pages where the forward button was already active. She said she did not notice. After identifying this question, I focused on whether other participants were pressing "*forward*" without selecting symptoms on some pages. Only P4 noticed it. Upon being asked the same question, P2 replied that he did not notice that either and that maybe the skip button would work. It made sense and I changed the prototype to add a "*skip*" button.

However, the last participant, P3, also objected to this idea. She said that she did not notice this because the skip button was the same color and position as the forward button, she had already read the list and decided to select some before seeing the skip button. She stated that it would make more sense to have a "*none of the above*" button in the list instead of a skip button. Since I found this data to make more sense, I decided to add the "*none of the above*" option to the required pages.

- P2 stated that as a user, he might want to see similar symptoms such as drooling, runny nose, and purulent discharge one after the other, as he had difficulty in finding the symptom of drooling. However, I did not find this idea feasible, as I did not see it possible to classify 130 symptoms in this way. For this reason, P2 stated that he would use the search box more often because he thought that he would use this application from time to time and would forget even if he learned the system.

- The symptoms that the participants had difficulty deciding were those related to tooth alignment and jaw relationships. I observed that the symptoms such as the sagging of the tooth and a deep bite or the symptoms related to the spaces were confusing the participants and the conditions such as the scissor bite or posterior crossbite were not understood. The reason for this situation may be due to the failure of the illustrations I made, or the inadequacy of two-dimensional visuals since orthodontic problems are three-dimensional problems. Nevertheless, it was meaningful in this respect that the two tested prototypes focused on the symptoms in this section. The results from here made me realize that as the case photos should vary according to the severity of the disease, visuals of healthy cases are needed for comparison, and three-dimensional models and animations are needed for some symptoms. In this way, I thought that misconceptions mentioned earlier could be avoided such as gingival recession as if it was perceived as a situation involving only more than one tooth.

- The unusualness of the opacity of the selected and unselected buttons (Figure 4.11) was only noticed by P2, the designer. Although this decision was unorthodox, I did not want to change it because I thought it would prevent the inactive buttons from mixing with the unselected buttons (Figure 4.13).

- During the interviews, I noticed minor problems with the interface prototypes. For example, I have observed that words describing the distinctive location should be bolded in some closely related symptom descriptions.

- I decided to exclude the symptoms related to jaw dimensions from the next algorithm version, as the users could not decide.

- After testing prototypes 1 and 2 with the participants, I asked which one they would prefer to use. While 4 of the participants preferred prototype 1, which regarding symptoms was divided into 6 steps, and took place on the same page, one participant (P3) said that she would prefer prototype 2 as it guided her while dealing with the symptoms, she is not familiar with. P4 and P5 expressed their opinions as follows:

> P4: I prefer the first prototype. The other one was too long for me. Long texts are difficult to read and understand. I can make an easier choice by looking at the pictures for myself. Here, for example, there is confusion. Just by looking at the picture, I can say whether there is or not. However other symptoms are difficult because some have very long explanations.
>
> P5: The first prototype gives more ideas and is faster.
>
> P4: There is a separate page for each symptom [in the second prototype]. However, in the first prototype, you listed all the symptoms one under the other. You choose very easily.

P2 stated, “It's like I would get bored when I get these questions one after the other. I wonder if there were 4 steps instead of 13?” After I said there were 23 steps, he said, “The shortest one if possible. A flow where options and images are in one place. For example, I took a minute, and it gave me a rough result. If it said if you give me 5 more minutes, I will show you more accurate results, I would be more motivated, I guess."

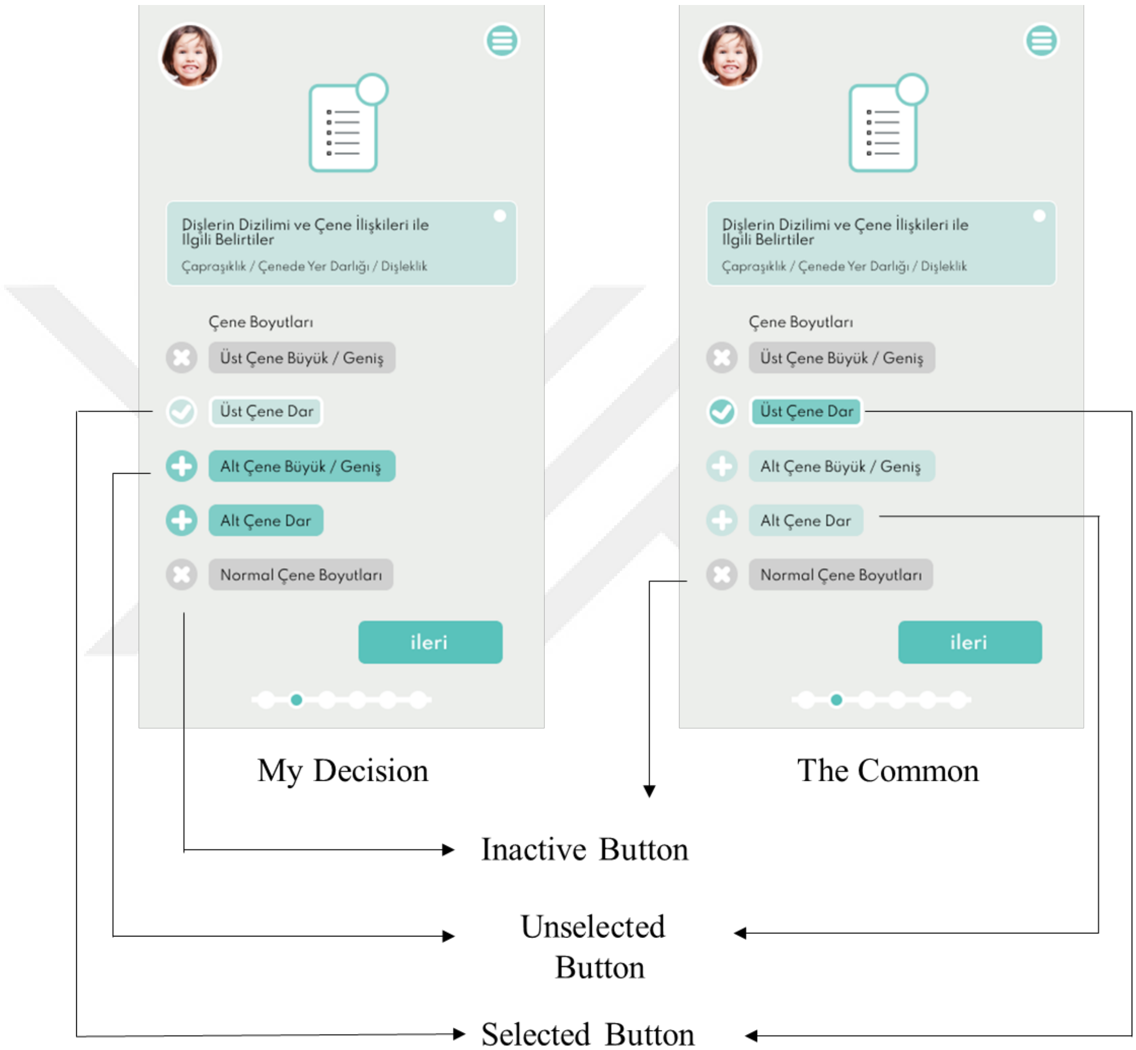


Figure 4.13: Comparison of my decision and the common point of view for the opacity of selected, unselected, and inactive buttons.

On the other hand, P3 stated:

> I think both are needed. If I already know what the symptoms are like pain, drooling, or fever, I don't need it to ask me whether it exists or not in every step. However, I do not know anything about oral and jaw health [orthodontic problems]. But here I am doing something about jaw health, AI will ask

something before I go to the doctor, and I will answer. That's why the second one felt more detailed and gave me more confidence… I think it's more helpful if it asks and guides me than if I try to find it myself. Otherwise, it felt like I had to check them all out. Asking me questions felt better in that sense. And for once you don't know, the subject is oral health. You are worried about whether I have given enough data. But since the prototype proceeds in the form of questions and answers, you will not feel this anxiety that much.

Based on these results, I opted to go on with the first prototype.

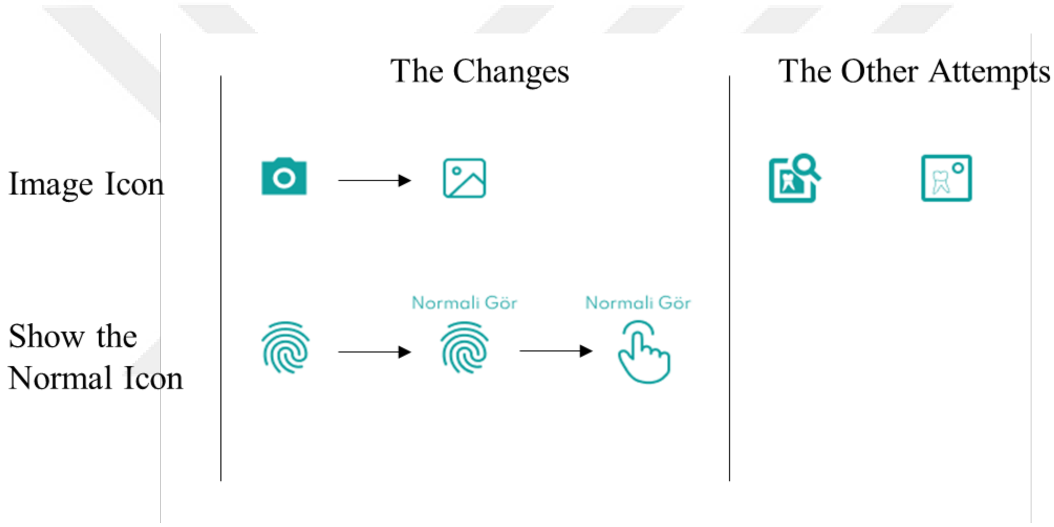


Figure 4.14: The changes of the icons over interviews.

The story that P4 expressed at the end of the interview was like explaining my purpose in designing this application. P4 stated:

> This app is important not only for children but also for adults. A dentist came to our company 20 years ago. He installed such cameras. They said the dentist came, for those who may wish to be examined, it was for free. I wanted to be too. By the way, I have no complaints about my teeth. The man is showing with the camera. It shows even the smallest blackness in my tooth. He told me that 12 fillings will be made. I was shocked. I said oh my god. He said I had 12 cavities that needed fillings to be done. How many teeth do I already have? After that, I was very sad and very scared, but I had no problems with my teeth. My teeth are fine. It is healthy. Then I said that there had to be something wrong. I went to

> another doctor. He also said that the previous dentist had exaggerated and that every blackness should not be filled. He did 4 fillings. Of course, after 12, I was ok with 4. Now the dentist we go to said, why were these teeth filled? I told him why. Then he said that such teeth should not be filled, there is nothing in your teeth. There may be some blackness. They remain for years, and you come and get them checked. If any progress is seen, then filling is done.

After I asked if the dentists had taken any radiographs, she told me that the last one had. This story was proof of the fact that patients need to know how the diagnosis is made so that they can choose the right treatment.

4.2.2.2 Preparation for the coding process

After the second research, I made the necessary arrangements in the interfaces, and then I started to prepare the prototype for coding. I postponed hiring the software developer for a month because the 3-month salary budget I had allocated for the software developer lost its value and I needed to improve the algorithm equation.

I tried to eliminate the diseases that should not be at the top of the results by adding new rules to the algorithm regarding the conditions they depend on. For example, the symptom of gingival bleeding can be seen both in gingivitis and when the tooth comes out of its socket (tooth avulsion). However, avulsion is conditioned by trauma. If there is no trauma, it should not appear on the results. So, by changing the equations and multipliers of some diseases in the algorithm, I increased the effect of the current conditions on the results.

While I was working on improving the algorithm, I was also looking for a software developer to code this prototype in the Android operating system. The biggest challenge I had when I found the right developer was getting them to sign a confidentiality agreement. Since I did not have an office where the developer could do the coding on the company's computer, I had to send the algorithm I developed to someone I met over the Internet. The software developers rightly did not want to sign an agreement that would be 10 times the salary they would receive after working for two months. As a solution, I decided to encrypt the algorithm. Thus, the programmer would be able to do the coding without knowing what the parameters in the algorithm were. So, I could hire a software developer without signing a non-disclosure agreement.

4.2.2.3 The coding process

While the developer was coding the backend of the application on Android Studio, I designed the interfaces and the architecture of the information that should be obtained in the child's profile. The 39 questions to be asked in the profile were classified under 4 main headings: General information, general health, nutritional habits, and parental health. I designed new pages by considering the flows in the scenarios in case the user has more than one child.

4.3 Testing Phase

The 2-month coding process was completed at the beginning of June 2022. I made tests to check the bugs in the application and prepared it for the performance test.

4.3.1 The third work package: Testing the prototype

In the project application, I had planned to test the prototype with real patient data. I designed a clinical trial to collect data from children who visited the clinic with a complaint. In the study, the data obtained in the anamnesis of the children and their complaints would be collected, and these data would be narrated anonymously and tested by two groups, the algorithm, and the pedodontists. I received ethical approval for this study from the Non-Invasive Clinical Research Ethics Committee of the Faculty of Health Sciences of Marmara University on 25.03.2021. After that, I decided to change the study design to include parents as a third group to investigate the app's usability. A power analysis was also needed to determine the number of participants and the sample size to reach statistical significance. Therefore, I received consultancy from Başkent University Statistics Consultancy and Research Center and applied to the same ethics committee again by adding the necessary documents and received my new ethics committee approval on 27.10.2021. However, since the prototype is diagnostic software, it is defined as a Class IIa medical device according to the MDR, and if a medical device does not have a CE mark, a human trial must get approval from the Turkish Medicines and Medical Devices Agency (TMMDA). While preparing the application, I noticed that there was an accepted ethical committee list on the agency's website and that the Marmara University Non-Invasive Clinical Research Ethics Committee was not on this list. When I contacted the committee and the agency, they said I needed approval from the ethics committees on the list.

As I thought that the whole application process would exceed the limited time of the project, I decided to have a pedodontist write the case stories. Thereupon, I applied to receive consultancy services from Başkent University, so that Prof. Dr. Ebru Tiralı could write case stories. After the bureaucratic processes were completed, she prepared the case stories by taking into account the diseases existing in the algorithm and the distribution table I sent her.

In the preliminary performance test, the case stories without diagnoses were evaluated by two groups. In the first group, I entered the data into the algorithm both in Excel and the app so that I could check if it was coded correctly. In the second group, the same case stories were evaluated by a pedodontist (Dr. Burcu Senirkentli). As a result, the algorithm showed 96% of the diseases in the case stories in outcomes, 93% in high-probability outcomes, and 80% at the top of the list. The pedodontist identified 83% of the diseases. The average number of diseases that appeared in the results was 14. The algorithm drew only one conclusion in some of the case stories and 30 in some others. The reason for the high standard deviation was that the algorithm was not sensitive enough and the case histories did not consist of isolated diseases.

Prof. Dr. Ebru Tiralı wrote her case stories in the form of stories involving more than one disease. The advantages and disadvantages of this situation were seen in the testing process.

The advantages:

- In real life, since the user cannot know which disease will be an independent disease, the entries to the algorithm may be like in these stories. Therefore, a realistic test result may have been obtained.

- In some of the complex case stories, the algorithm infers a high number of diseases. These results will be leveraged to develop new elimination techniques in the next version of the algorithm.

The disadvantage:

- Secondary diseases turned into noise in the results because the data in some stories were not sufficient. This has led to a large number of results and secondary diseases to be listed below.

Real performance can be achieved by parents entering their own children's data into the algorithm. Parents who visit the pedodontics clinic in a university hospital will be asked to enter the complaints they see in their children into the algorithm; The result obtained will be compared with the clinical diagnosis of the child. This research will be one in which researchers from two fields such as dentistry and design work together, and the results are evaluated from two perspectives.

Before starting to describe the third research, it is necessary to mention the achievements in this period. In September 2022, I participated as the founder of one of the 11 companies that were invited to the "Dialogue with Key Stakeholders on Universal Access to Digital Oral Health" event organized by the University of Montpellier. I presented the project to public health experts, academicians, and the chief dental officers from WHO. In December 2022, I was one of the 58 scientists who got accepted to the "Leaders in Innovation Fellowships Program" of the Royal Academy of Engineering.

4.3.1.1 Research 3: Evaluation of the application with users from 3 different fields of expertise

Since I was going to submit the last report for TÜBİTAK in March 2023 and I was already supposed to test the coded prototype with users, I wanted to do a more comprehensive study and include it both in my thesis and the report.

In this research, I decided to include 9 participants from 3 different fields of expertise, 3 dentists, 3 designers, and 3 parents as non-experts to evaluate the app (Figure 4.15). My primary aim was to understand the experiences of the participants from three different backgrounds that make up the pillars of this project. Of course, parents were the target users of the app, but I was also aiming to collect data from dentists on the medical content and designers on the design of the interfaces. Like the previous user research, my goal was to improve the design of the product.

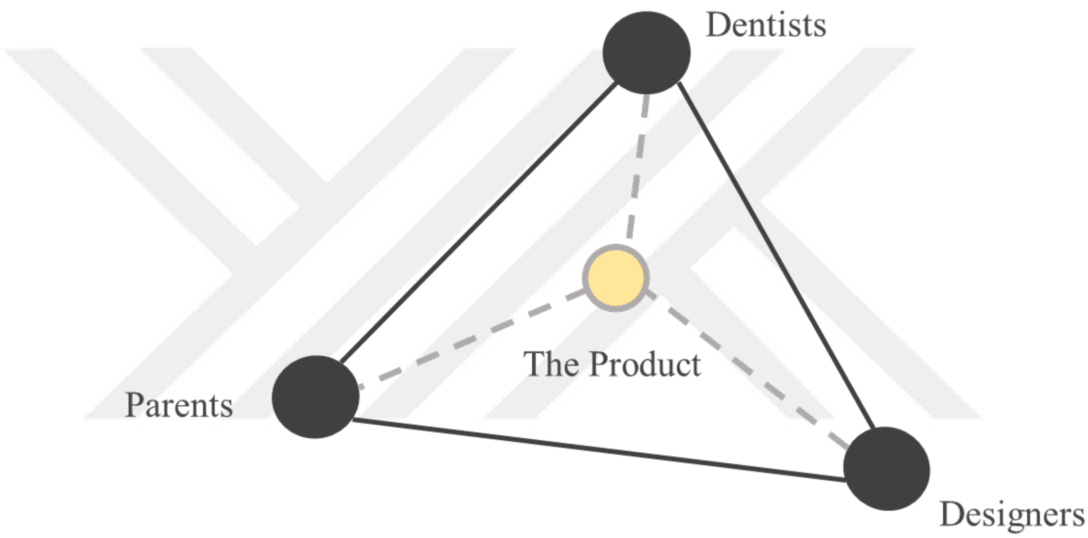


Figure 4.15: Three different poles according to the backgrounds of the participants.

As a researcher, I also wondered how different the experiences of participants using the same app and coming from different backgrounds might be. What did they mostly focus on and pay attention to while using it? In research, you do not ask the question you already expect the answer to. It's not unexpected that dentists talk about medical content or designers talk about interfaces but what is the distribution of these points in the interviews? Can this method be valid for designers? Answering these questions was my second goal.

I wrote a fictitious case story of a boy whom I named Ali. I decided to ask participants to complete two tasks: Create Ali's profile and make an assessment based on his health information and complaints in the story and think aloud during the assessment. In order not to bore the participant during the evaluation, I did not specify the answers in the story to all 39 questions in the profile, and since I would not compare the accuracy of the results, I said that they could answer the questions that did not have an answer in the story as they wished. The information in the story for the profile is shown in Figure 4.16.

Name: Ali

Date of Birth: 16.02.2016

Weight: 40

Height: 1.2m

Systemic Disease: Reflux

Congenital/Hereditary Disease: Lip Palate Cleft

Allergy: Dust

Medication: Antihistamines

Trauma History: Face injury in 28.11.2022

Ali has been drinking soda every day since he started school. He likes sugary foods.

Figure 4.16: Ali’s health information given to the participants.

I determined three different situations describing the symptoms they saw in Ali for the evaluation they would make after filling out the profile. Each situation corresponded to a different section in the symptom list. I decided to give only one of these situations to each participant with the same expertise to not bore them by making the interview too long. It was enough for me that they experienced the evaluation process in general. Figure 4.17 describes three situations in which each participant evaluated only one.

I received ethical approval from Istanbul Technical University Health and Engineering Sciences Human Research Ethics Committee on 24.11.2022 for this user research. I included nine participants in total (Table 4.2). I did the interviews via Zoom by recording audio and video. The participants used the prototype via the emulator on Android Studio. I transcribed the 9-hour-long recordings thematically and sometimes I also took notes of things they did while using but didn't say out loud. I also evaluated the comments based on six main categories: thematic ones, those related to medical content, those related to interface and user experience, those related to technological problems, and general comments and notes I took. I made a color-coded map out of these results.

Situation 1:

You think Ali has been neglecting brushing for a long time. He had bleeding from his gums when he bit down on the fruit a day or two ago. You see him spitting blood when you tell him to brush his teeth. When you looked into his mouth, you saw widespread redness and swelling of his gums. You see a yellow deposit on the teeth.

Situation 2:

You noticed an abnormality in Ali's jaw closing. The upper jaw is narrow and the lower jaw seems forward. In fact, the upper teeth are positioned behind the lower teeth, as if the lower front teeth are covering the upper front teeth. Because Ali has allergies, his nose is often stuffy and he sometimes snores at night.

Situation 3:

Ali says he has been having pain in his teeth lately when eating something cold. You see transparency or wear when you look at his teeth, especially in the back teeth. The upper second right incisor is small and the left one never came out.

Figure 4.17: Three different situations Ali experiences.

As I stated in the research approach, I evaluated the results in two categories according to the purpose of the research. While I discussed the results, I obtained to improve the use and the design of the product within the scope of "*Research for Design*", the results I obtained in response to the questions for my secondary goal took place within the scope of “*Research into Design*”. Because discussing the distribution of comments and making color-coded maps were beyond the aim of RfD. At that point, I became a researcher willing to contribute to the academy and the product turned into just a means.

There was a total of 254 comments including my observations (N:35). To categorize the comments, I prepared two Excel files: In the first, I gathered the comments made by each participant under 6 categories: thematic ones, those related to medical content, those related to interface and user experience, those related to technological problems, general comments and notes I took. To categorize the comments, I prepared two Excel files: In the first, I gathered the comments made by each participant under 6 categories:

thematic ones (the comments made by the participants by making inferences on Ali's data), those related to medical content, those related to interface and user experience, those related to technological problems, general comments, and my observations. In the second, I reclassified the comments according to their content. For example, I assessed the comments on medical content in 4 classes: comments where the question is not understood or thought to be incomprehensible, confusion about to which part of the profile the data fits, comments on the ambiguity of answers, and general comments.

Table 4.2: Backgrounds of the participants included in research 3.

Participants	Age	Background
Parent 1	38	She is a mom of a 3-year-old. She works as a lecturer in the language department at a university.
Parent 2	38	She is a mom of a 2-year-old. She works in a retail company as a manager.
Parent 3	37	She is a mom of a 2-year-old. She works as a lecturer in the physics department at a university.
Dentist 1	47	She is a professor in the periodontology department at a university. She is a mom of a 19-year-old.
Dentist 2	38	He works as a clinician in a private clinic.
Dentist 3	38	She works as a clinician in a state hospital.
Designer 1	35	He works as a design lead in a gaming company.
Designer 2	43	He works as a lecturer in the industrial design department at a university
Designer 3	32	He works as a freelancer designer in the fashion industry. His field of interest is brand management, website, and logo design.

While evaluating the results in terms of RfD, instead of concentrating on what the participant from which background said, I focused on how the participants encountered problems in their interaction with the product, the comments about the content and interface, and the changes I will make in the product by looking at these comments. While evaluating the comments for RiD, I visualized which participant group focused on which topic during the research process on a colored map and discussed how this method enriched the data.

4.3.1.1.1 The results of research 3 from the RfD perspective

A total of 79 comments were made on medical content. Rather than listing the comments made for each section in the app, I find it more convenient for easy reading to summarize which design decisions I made based on which type of comments. The decisions I made, and the examples of comments are as follows:

1. The majority of the comments (N:53) were related to poor understanding of the medical terms in the questions or concerns about not being understood by the laypeople (these comments were mostly made by dentists). Therefore, it is necessary to rearrange the questions in the profile. What needs to be done for these questions can be grouped into four classes.

    a) First, if a term must be used in the questions, it is necessary to put an explanation in a way that lay people can understand. For example, the participants commented on several terms as follows:

        Parent 2: I have never heard of a test called chorionic villus sampling.

        Designer 1: Does every parent know what infectious means? An explanation may be needed here.

        Designer 2: What does bedtime mean? The time during sleep or before sleep?

        Designer 3: What is BMI? There may be an info sign next to it for further information.

    I also noticed that some general terms in literature do not correspond to the examples in real life. For example, Parent 3 did not understand how to answer the question about the duration of bottle use. She asked “Is this just a bottle? Aren't things like drinkers or pacifiers included? If the pacifier is included, I will say yes, but if it is only a bottle, I will say no.” In such questions, if there is no exact answer in the literature, it may be necessary to design the question taking into account the underlying factor. It is reasonable to include the pacifier in the bottle question since long-term sucking has a negative effect on the child's jaw relationships.

b) Some questions should be merged. For example, I designed the questions in the profile about the child's general health by considering the questions I asked while taking anamnesis from the patient in the clinic. During the anamnesis, I used to ask questions by classifying each type of disease (Do you have a systemic disease such as heart, diabetes, or blood pressure? or do you have a hormonal disorder?). Therefore, I compartmentalized the diseases and disorders as systemic, genetic, congenital, hormonal, contagious, infectious/early childhood, and psychological. However, after this research, I noticed that users cannot make this distinction. Some participants did not understand what was meant by systemic disease. The comments were as follows:

> Parent 1: I guess reflux is not considered a systemic disease. Is reflux a systemic disease?
>
> Parent 2: What does a systemic disease mean? I do not know. Is reflux a systemic disease?
>
> Designer 2: What do you mean by systemic? Here the question does not speak the same language as me.

The dentists in particular discussed which diseases the systemic disease represented exactly and made recommendations:

Dentist 1: For example, can an allergy be considered systemic?... Instead of systemic diseases, you can write it as a condition or disease related to your general health.

Dentist 2: Is reflux a systemic disease? Is knee pain systemic? It is not systemic; it is an orthopedic disease. Systemic disease may be anemia. What is your child's illness? It was a more correct question.

Dentist 3: For the systemic disease part, it would make more sense if the question was something like Does your child have any disease?

Thereupon, I decided to gather these questions under one title and let the algorithm make this distinction itself because it is unfair to expect users to distinguish the diseases according to this category.

Another intervention of dentists was to the question "*Do you or your partner have active caries?*". One of the things I often encountered in my conversations in my social life was that people did not fully understand the difference between caries and a filling. I met a lot of people who told me they had a lot of caries in their mouths and showed me their fillings. Therefore, I wanted to emphasize "*active*" in the question. However, two dentists commented on this question:

Dentist 1: They may not understand what active caries means. You want to distinguish untreated caries from filling, but you need to express it in another way.

Dentist 3: They may not understand active caries. Do you currently have untreated decayed teeth in your mouth? It is clearer.

c) During the R&D process, I did not have time to design the medication part of the profile with trade names. For this reason, I created a list of drugs with pharmacological and generic names mentioned in the literature and I was aware that users would not understand any of them. The participants also frequently expressed this in their comments:

> Designer 2: When a medicine is asked, I would normally look for the name of the drug.
>
> Parent 1: Can you tell me a trade name for the tetracycline? I would like to learn.
>
> Parent 2: I don't know what thalidomide means. Like what is this?
>
> Parent 3: Is tetracycline a component of a drug? But only a careful parent might know. The doctor prescribes it anyway, I probably wouldn't have known, and what if I don't know? There is no such option here. Also, I think that you are taking the child to the pediatrician, and he prescribes an appropriate drug, so I do not question it.
>
> Designer 1: What is tetracycline? What is derivative? Presumably, both are a type of antibiotic. Imagine, I am asking my wife: Did our child use a tetracycline derivative antibiotic?... What?... Derivative, derivative… (He laughs) A colloquial explanation is needed here.
>
> Designer 2: Tetracycline derivative antibiotic is very technical.

It is necessary to create a list of licensed drugs with their trade names for each country. Maybe a barcode reading system can be added for ease of use.

d) While preparing the algorithm, I took into account all the etiological risk factors in the literature and did not question how valid these parameters were. Because of the comments below, I've decided to re-evaluate the validity of some questions in the profile and consider whether they should be asked to everyone.

For example, I decided to remove the question of Thalidomide drug use from the system, which was released in the 50s for morning sickness in pregnant women and caused advanced birth defects in children. Although this drug is currently used in cancer patients, I realized that the negative impact on the user experience of asking this question is more than the contribution of the answer.

I was considering removing the tetracycline question as well. However, Dentist 3 reminded me of the wrong use of antibiotics in Turkey and made another suggestion:

> Dentist 3: In which diseases is tetracycline used? Parents may not know. It may be necessary to write the trade names."
>
> Me: Yes, I know. I added these parameters because they are mentioned in the literature, but doctors may not prescribe tetracycline to pregnant women and children anymore.
>
> Dentist 3: But maybe she has it at home, she may have used it according to her mind… It might make more sense to ask the tetracycline question only to people with discoloration-related symptoms rather than to everyone. Because it will probably be very rare.

After Dentist 3's comments, I decided to convert the question about the use of tetracycline-derived antibiotics into a question triggered by the symptom rather than a regular one in the profile.

2. Ten comments were made about not knowing where the information in Ali's story belongs on the profile and the purpose of the questions. The results are as follows:

    a) Some participants searched for the lip-palate cleft in the systemic disease section. Gathering all diseases under one title can solve this problem.

    b) Two parents asked whether lip-palate cleft causes disability or not. This can be considered the limitation of the study. I believe that if their child had this defect, they would probably know the answer.

c) There was an issue that I didn't think of while I was preparing Ali's story, but the dentists noticed. Postnatal surgery is performed for children born with a lip-palate cleft and the defect is usually corrected. This causes trauma to the face area. While entering Ali’s trauma history, Dentist 2 asked, “A surgery is performed for the cleft after the birth, will I add it here?” This question showed me that the algorithm should automatically add the trauma parameter to the database without the user having to think about it.

d) Since the algorithm cannot understand texts and traumas can be very diverse, the system only asks for trauma dates. The comments were as follows:

> Designer 3: It's asking about trauma dates here, shouldn't it also be asking what trauma is?
>
> Designer 2: It asks about the child's trauma, but it doesn't want the story, it asks for the dates. This felt a little strange. Because it's hard to remember. It wants the full date. At that moment, this information in the forms jams the user. This information is not available at the moment, I would like to continue when I find out about it. I don't want to continue without answering it.

I noted that this should be worked on, but I did not make a design decision.

e) In the treatment section, I had added the ones mentioned in the literature that only affect the child's oral health because it was not possible to list all treatments. Therefore, I had also added a box at the end of the list where users could add the treatment as text for UX purposes even though the algorithm was unable to understand such entries. Dentist 3 stated “Hmm, our child had reflux, but it is not on this list. Then I'll add a stomach protector and skip it because he has reflux.” On the other hand, Dentist 2 criticized this section:

> Dentist 2: Why is there a text box at the bottom of the list?

Me: Treatments can be very diverse, so I'm taking it as text that the algorithm won't understand but will record.

Dentist 2: It may make sense to remove some things. You don't want to overlook the majors just because you want to get all the information. Are treatments that important?

Me: The probability of oral mucositis is highest in children receiving chemotherapy or radiotherapy. I don't want to skip this one either.

Dentist 2: Don't parents whose children have chemotherapy know this?

Me: But I didn't edit the algorithm accordingly. In addition, doctors do not understand much about oral diseases. The algorithm can also be used for doctors.

Dentist 2: Right, I still see it as something I can use in the clinic.

After this conversation, I decided that the question about treatments should be conditional and triggered after the selection of relevant diseases.

3. The reason why some questions are unclear is because the parameters mentioned in the literature are not precisely defined. Seven comments were made about the vagueness of the questions:

   a) One of the causes of enamel defects seen in children in the literature is the exposure of the mother to environmental pollution. Therefore, the algorithm asks the users in the parental health section whether they have been exposed to environmental pollution during pregnancy. The comments on this issue were as follows:

      Parent 1: Am I exposed to environmental pollution? Well, if I lived in China, I would say yes, but are we being exposed? I said no, I don't know whether.

Parent 2: Being exposed to environmental pollution is a very broad question. For example, I am possibly exposed to environmental pollution, but I do not live in Zonguldak. I will still say yes.

Dentist 1: How should we answer the question of environmental pollution? Air pollution may be less in Ankara, but it may be more in the surrounding areas.

Dentist 2: Is there environmental pollution? This question is very broad. A person living in İzmir is likely exposed to environmental pollution.

Designer 2: They may not perceive air pollution as environmental pollution. Do I see garbage around? No, I don't see it, so they can say that there is no environmental pollution.

At this point, it may be more appropriate to ask about the location where the user lived while pregnant and to use a database on environmental pollution, if exists.

b) Parent 2 commented on the question of whether the child frequently consumes acidic foods such as pickles or lemons: "What does frequently consumed mean? It needs to be specific." Parent 2 is right, however, the answer to this question is not found in the literature. As a solution, it might make sense to convert this question into a triggered one.

c) Another confusion in the literature was related to the duration of breastfeeding. While pediatricians recommend that children suckle for at least two years, pediatric dentists say that children should not suckle for more than one year because it causes caries. When children suckle for less than a year, their maxillofacial development is retarded. Therefore, I had to determine the answers to the question about the duration of sucking, as less than 12 months, 12 months, and more than 12 months. Designer 1 commented on this question:

> “How many months has your child been breastfed? Times seem strange to me. It may be technically correct, but practically incorrect. If this were 364 days, wouldn't it be 12 months? I understand that this is the case in the literature, but I would advise you not to rely on the data you get in this way. If you can't get reliable data, don’t ask this question.”

As a solution, instead of restricting the answers to this question, users may be asked to type a number themselves and the algorithm can calculate the risk according to the typed duration.

4. Some comments were those that the information gathered in the profile did not match the target user who will use this application. The comments were as follows:

> Parent 1: I wonder if a parent who constantly gives candy to their child uses this application.
>
> Dentist 1: I don't think there is anyone in Turkey who would say no to the question of stressful events in the family.
>
> Dentist 2: I understand that you want to encompass everything, but would a parent whose child smokes use this app? The questions need some refinement.

Since the algorithm cannot be arranged with generalizations about the lives of the users, I did not make a decision according to these comments.

5. Dentist 3 strongly criticized the symptoms related to the alignment of the teeth and jaw relationship, while entering the data of the third situation. After reading the situation, she said that parents might not know this problem was orthodontic. She stated, “I couldn't even remember what concave or convex was... Backstep in canines and molars, what? (She laughs) I think the parents of children with orthodontic problems bring them to the clinic because of the crowding. It is okay

for dentists, but this information is too much for parents… Decreased lower jaw height? I think that's too much, even for me."

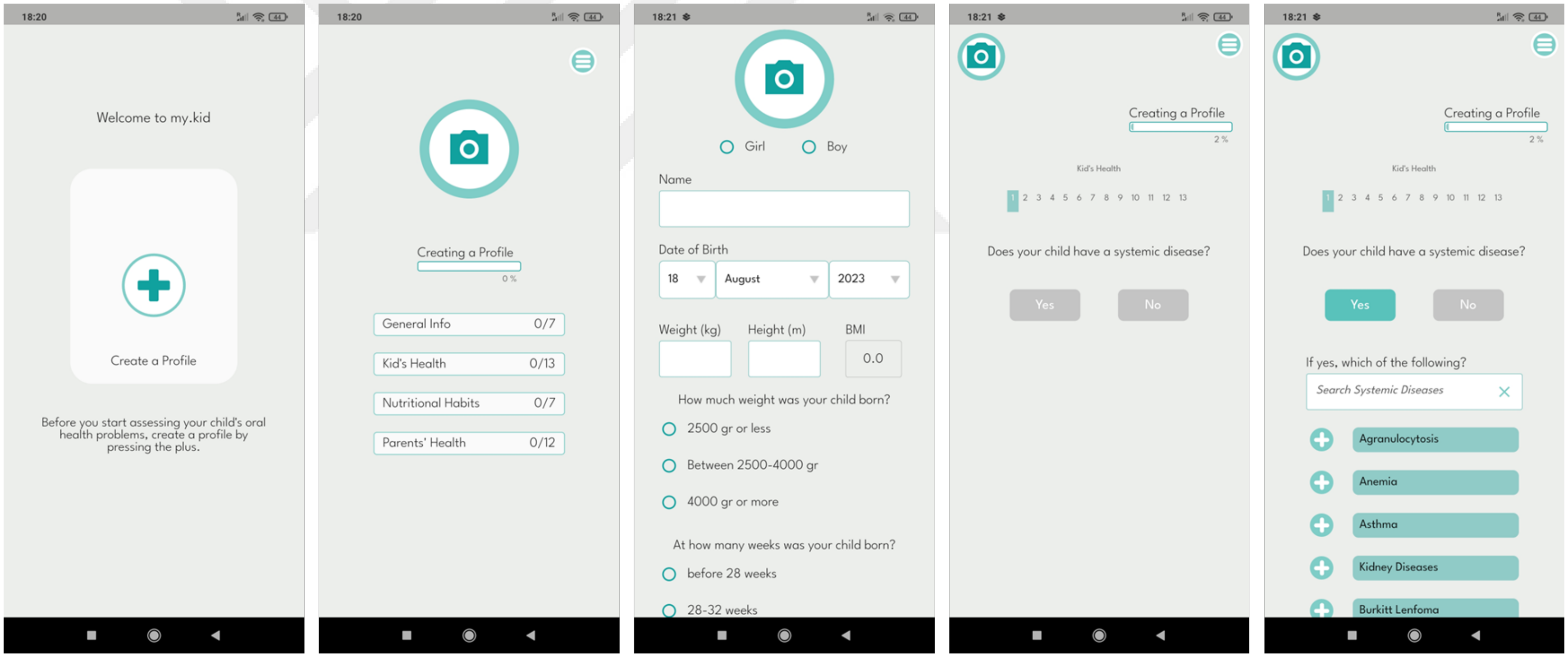
I still think that with a few edits and good visuals, this part could be used to simulate a child's mouth. Because in the second study, the participants did not have great difficulties in assessing these symptoms.

A total of 121 comments were made on the design of the product. I think the reason why this number is so high is that the interfaces were designed by an inexperienced and untrained person like me. If the interfaces had been designed by a professional, I do not think that this number would be that high.

I can say that I had a very difficult time categorizing the comments about the design and I could not make an evaluation based on a certain criterion. First, I tried to categorize them as those related to product communication, UI elements, user experience, navigation, functional features, and human factors. However, these concepts are intertwined; unfortunately, this categorization did not work. This is because it was not possible to fully separate UI and UX and how the interface elements are designed directly affects the user experience. Then, I decided to cluster the comments according to the pages on which they were made and some specific design elements. Although I had the urge to unfold every deduction I made, I only list the comments that I think stand out.

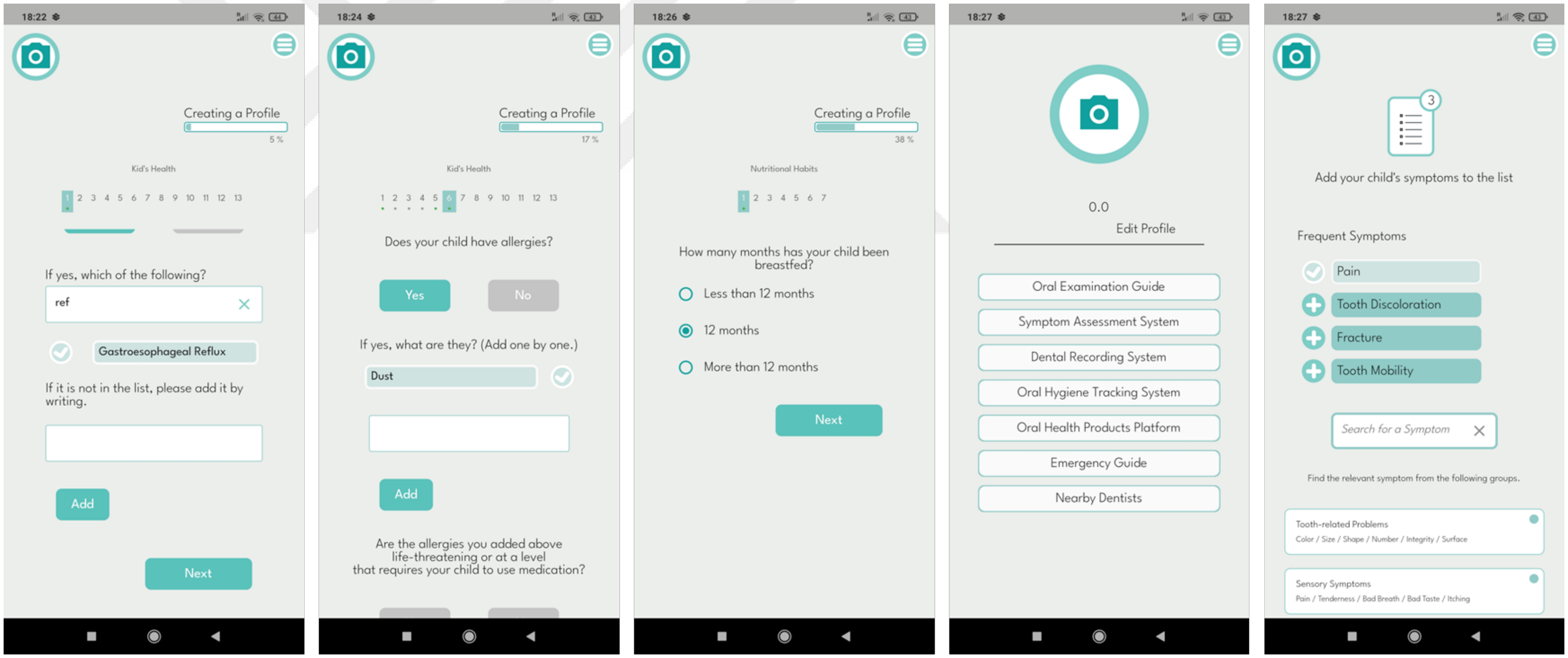
1. The issues in the creating profile can be summarized as follows:

   a) It is very important to create a profile of the child so that the algorithm can draw more precise results. For this reason, when the user downloads the application for the first time, there is a Create Profile button on the landing page. The comments on this page (Figure 4.18.a and b) were made as follows:



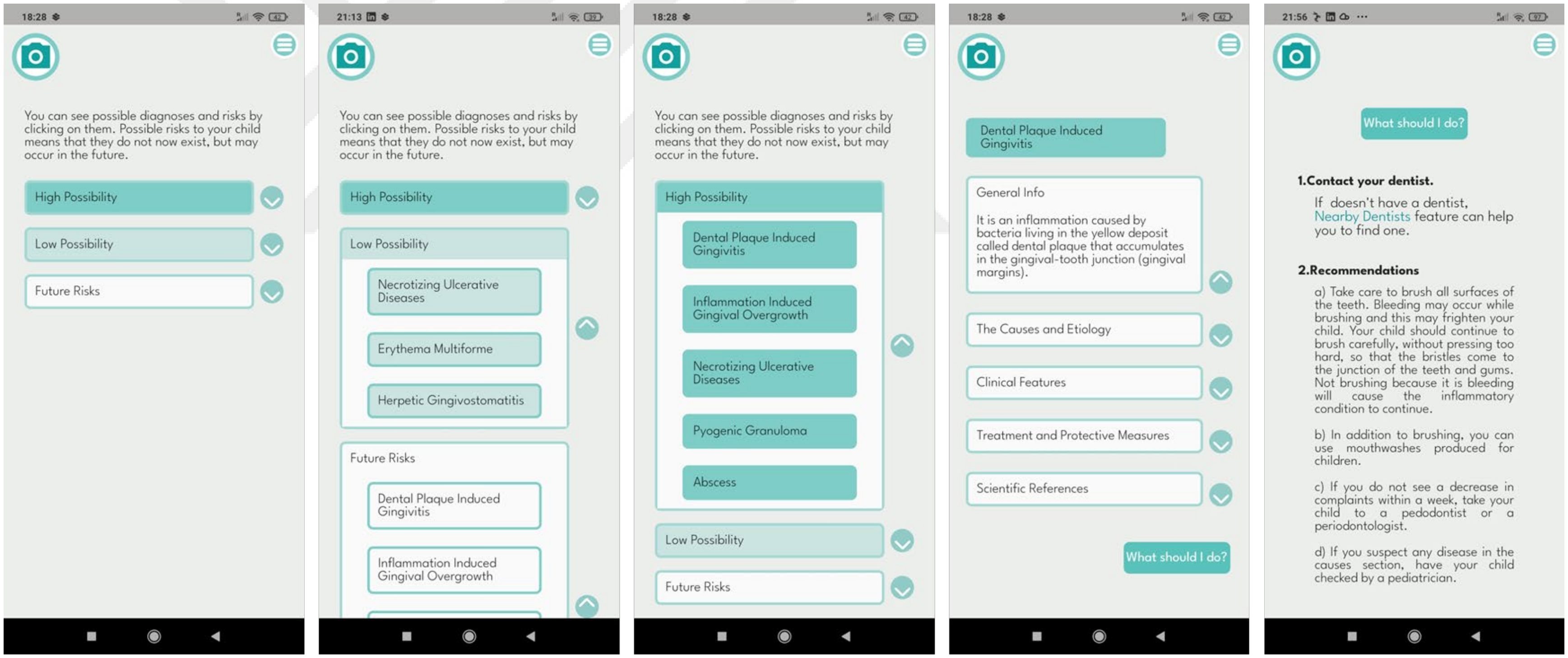
(a) (b) (c) (d) (e)

Figure 4.18: Screenshots of interfaces (Part 1/3).



(a) (b) (c) (d) (e)

Figure 4.19: Screenshots of interfaces (Part 2/3).



(a) (b) (c) (d) (e)

Figure 4.20: Screenshots of interfaces (Part 3/3).

Dentist 2: The landing page is very quiet.

Designer 1: We say create a profile; this will probably work but it is a different call to action than my purpose of installing this application. I came to evaluate my child's oral health. After reading this long paragraph, I understand that I need to create a profile first. It would be better if a sign such as Create a Profile then Assess was given to me so that I could have seen the next step.

Designer 3: There's the Create a Profile button, it's actually nice that it's bold. It feels like it's medical. On the other hand, if there was an icon at the top, it would be nice to see how many profiles I could add. There are extra profiles on Netflix’s starting page, it would be nice to see the empty slots.

During the design phase, I thought about putting empty slots on the landing page and I designed a few prototypes, but later, I didn't want the user to be faced with a crowded page at every launch if they have only a child. Instead, I found it more appropriate for the user to use the "add sibling" feature in the side menu to add a new profile. As Designer 1 said, the main motivation for users to download the app is to make an assessment, and creating the profile tab may not be something they expect. For this reason, elements that will show the evaluation process to the user with every step must be added to the interface.

b) It is necessary to rearrange the information architecture and hierarchy in the profile interfaces. For this, it is necessary to make arrangements to make it clear to whom the profile belongs, to separate the information on the page with cards according to their relations with each other, to add icons for specific functions, and to arrange the spaces between the components according to this hierarchy. Some of the comments associated with this issue (Figure 4.18.c) were as follows:

For the gender section:

Designer 1: We see the name and the date of birth as explanations on the page. We should also see the gender.

Designer 3: The girl and boy part escaped my notice. I focused directly on the name. I wondered if those gender tabs should be there as a position.

For the photo section:

Designer 1: We are creating a profile; whose name and photo will I enter here? The only sign showing me that it won't be me is that it says girl here. You need to put an expression like your child's profile… Clicking on the photo opens the camera. It might make sense to put a camera icon at the bottom right of the photo.

For the birth weight section:

Designer 1: The distance between the question of with how much weight your child was born, and the previous component should be a little more. To give the feeling that one question is over, and another question is starting. It might make sense to separate the questions with cards here.

c) There was a row with numbers at the top of the page to show the 13 questions in the section about the child's general health. This bar was inadequate in terms of navigation and gave some users the wrong impression. Three users mistook the top bar on the pages asking questions about the child's current illness to be a Likert scale (Figure 4.18.d). The comments were as follows:

Parent 1: I thought that it was a Likert scale.

Parent 2: I thought it was asking how I found my child's health. I rated it out of 10… I understand the system now. For example, the dots under the numbers show the questions I answered.

Dentist 3: I thought that it was a scale.

Designers' opinion on that bar:

Designer 1: If it is something that goes left and right with a swipe, these numbers should be at the bottom, not the top.

Designer 3: Actually, it's nice to see the question steps, but it seems like it was just put in without thinking.

d) Four comments showed the issue where the keyboard covered the search results obtained via the search box. Results may appear at the top in search boxes and spaces on the page can be used more effectively.

e) Dentist 3 and Designer 1 had opinions on the listing of all diseases (Figure 4.18.e):

Dentist 3: Listing diseases is pretty good because it doesn't come to mind right away. Patients perceive chronic diseases as a part of normal life. For example, when I take anamnesis, I ask the patient if he has a heart condition, diabetes, or high blood pressure, and he says no. Then when I ask him if he is using a blood thinner, aspirin-derived medicine, he says yes. When I ask him why, he says he had a stent. The man does not know that he has a heart disease.

Designer 1: It may not be necessary to show all diseases. Or it is more logical to make a drop-down menu with a phrase such as see as a list. It would be better to go to a separate page for that, by the way.

2. Problems encountered in the symptom list pages during the assessment process are as follows:

a) I noticed that the frequent symptoms section positioned at the top of the page misled the user more than they should (Figure 4.19.e). The participants had the impression that they would initiate the assessment based on the symptoms there. Many of the participants chose symptoms from this section, even though they were not relevant to the data in the story. The comments were as follows:

   Dentist 1: When I saw the frequent symptom section, I realized that I was trying to fit the symptoms in the story to the ones here. For example, I selected discoloration here due to the symptom called translucency in the story, also pain instead of sensitivity. Maybe it would make sense to position the search box at the top.

   Designer 2: I see frequent symptoms first. Here I choose a symptom. It feels like I must pick something from this section. I scroll down and the thing I selected disappears from the page. Then, I have to choose something from the list below. I'm torn here. That's why I said I didn't know where I was. It created the perception that if I had chosen something else in the list above, other things would appear in the list below.

b) While evaluating the symptoms, I noticed that no one pressed the picture icon (Figure 4.14), and I asked the participants the reason why. The comments were as follows:

   Dentist 3: Well, it never caught my attention. It felt like it was just standing there, and when I pressed on it, it didn't feel like it was going to explain something to me.

   Me: If there was a sign, like info, would you be interested then? Or what would make you think that you would see an illustration here?

   Dentist 3: Maybe you can place an explanation, but of course, this is my experience, maybe I don't press it… It doesn't say press on me.

Designer 3: Is it asking me for a photo here?

Me: Did that icon give you the impression of uploading a photo?

Designer 3: It didn't, but it didn't make the right impression either.

Dentist 1: A colorful picture icon may encourage users to press on it.

c) The comments on inactive buttons (Figure 4.13) showed that this method did not direct the user as I had planned. The comments were as follows:

Dentist 3: Why did this happen? X's have appeared.

Designer 3: Why are these crossed out?... It makes me feel like I guessed it right. It is an interesting experience. I don't think this impression is correct. At least it shouldn't be a cross.

3. The problems seen throughout the application are as follows:

a) When participants chose a disease in the profile or a symptom, they were unsure whether they had chosen it. While some of the participants accidentally chose an irrelevant option and did not realize it, some of them made the right choice but felt the need to press again the Add button under the box that was placed to add an option that was not on the list (Figure 4.19.a). A total of 23 comments were made regarding this issue. Some of these were as follows:

Parent 2: Actually, there was a tick next to it, but I couldn't add a symptom a couple of times before and I felt the need to press add because of it. I see the add button first, but the next button on the page later, I guess that's why I clicked on the add button. Maybe if the two line up I wouldn't press it.

Parent 3: Am I choosing the right one by pressing next now?

Dentist 1: I chose it, is there any need to add it again? I guess not.

Dentist 2: Did it save the things I selected now?

Designer 1: Does your child have a contagious disease? No… No button went unchecked. I couldn't understand what I marked.

Designer 2: The add button made me feel like I had to press it to add it like add to cart. It felt like it wouldn't be saved if I didn't click it.

Designer 3: Where is the thing I chose? Is it saved?... I wish I knew what I added. It would help if it appeared somewhere when I added the drugs.

These comments showed that the buttons and the selection action had poor communication with the user. I decided that the selections made should be seen somewhere on the page or it should be possible to see the profile created at the end of the process and motions that will emphasize the selecting action should be added.

b) Participants had difficulty navigating between sections and browsing lists. Part of this difficulty was related to the fact that they used the app using a mouse on the computer during the research process. Aside from this limitation, they struggled, especially when scrolling down long lists of diseases and scrolling through sections in the profile. When I evaluated 13 comments on this issue, I realized that there is a need for markers in the interfaces that show the user's location and what stage they are at in the whole process. Some of them were as follows:

Dentist: 2: There are so many buttons on the homepage, and it scared me a little. My purpose is to evaluate. Now what are the oral examination guide and the emergency guide? Or am I going to go to the dental record? (Figure 4.19.d)

Designer 1: Does your child have a systemic disease? I am lost now. If I press next, I should see some progress. It's like on e-commerce sites. The profile completion percentage didn't mean much. Looks like there should be a navigation bar here… How do I get to the bottom of the list? It is necessary to scroll down the whole list. (Figure 4.18.d)

Designer 3: Have I switched to eating habits now? I saw new steps. This can be a little confusing. Its hierarchy is not fully understood… It's like I forgot where I was and what I was doing. I was creating the boy's profile. What am I doing now? (Figure 4.19.c)

4. On the results page, different comments were made regarding the display of the results and the experiences of the participants.

   a) The comments on the display of the results were as follows (Figure 4.20.a,b,c,d,e):

      Parent 3: Is there any probability of the results?

      Dentist 1: Are these results in order of risk or do they have the same hierarchy? Maybe they can be colorized.

      Designer 1: There should be something like read more, not a drop-down menu here. I read a few lines and if I want to read more, I have to press on it.

      Designer 2: Is there any point in giving high and low-probability information? I think there is no need to split the results this way and open them in accordion form. I prefer a results page that I can read in one go. It's the first thing I click when you say high, and I'm disconnected from other options. It should not make me neglect the others.

      Designer 3: There is a box on the result page, I click, and another box opens. It makes sense, but it doesn't feel like I'm going to click on it. It may be because of the color. Maybe it needs a call-to-action, maybe something like learn more or read more, you know, it's underlined. It could be something like this. Why are the boxes at the bottom as an outline and these are full?

   b) The comments on their experiences were as follows:

Parent 1: I was very impressed with the results. I think it diagnosed it pretty well. I don't know if I would look at all of them one by one, but I would look at the high probability ones. It's also nice to refer me to the dentist. Sometimes you go to the doctor and your problem is not solved. Maybe you can research here and focus on low-probability issues. I really liked it. I would keep this on my phone. It would not be an app that I downloaded and then deleted after the use.

Parent 2: Overall good by the way. I would be more comfortable if I used it on the phone. I didn't understand the first question. But on the other hand, I think it's easy. If some questions are getting specific, then it would be much better. Having a photo is very helpful. People understand better what they see. If it were me, I would judge by looking at the photo. Other than that, it's pretty comfortable.

Parent 3: Obviously, this product is usable. People expect to receive information according to their situation. For oral diseases, you feel like you must go to the dentist because you cannot google something visual and specific. It's valuable in that respect. It also actually makes sense to record and track information. Things like routine checkups can also be put in. I didn't need it until now because my kid is small.

Dentist 1: There is so much going on in the results. It's as if we put the perfect doctor in the system here, it draws a lot of results, but we leave the judgment about what the result will be to the user. This would be unfair to the user. It would be better if these results were shown to physicians…Yes, you are right about the legal responsibilities. Let's just say it drew a conclusion. The user received information about erosion and went to the doctor. The doctor looked at it and said it was tooth decay. Then what will the user say? No, this is an erosion? It makes sense to have several possibilities here.

Dentist 3: Seeing a lot on the potential risks page can make parents nervous, especially in Turkey.

Designer 1: It said so many things to me that it seems like it didn't say anything… The process took too long for me. Of course, I understand that it asks for that much information to make the correct diagnosis, but the process is still long, and, in the end, it tells me to go to the dentist. I'm using this app to not go.

Designer 2: I click on What should I do… (He's laughing) This answer is the thing that annoys me the most. What should I do? You should contact the dentist. It makes me feel like I can do something, and it finally leads me to something I already know I must do… Instead of What should I do, something else can be written.

Designer 3: "Interfaces feel drafty. But the content makes you feel like you're encountering something competent."

Looking at these results, I can summarize my decisions as follows:

- Medical content should be converted in a way that the user can understand. For this, vagueness in questions and answers should be eliminated, and medical terms should be used as descriptively as possible.

- The communication of the interfaces should be reworked. User navigation should be facilitated. Interfaces should be redesigned by a professional. I believe that the results of this research will be useful in the next stage.

- Eliminating some questions and shortening the profile-creating process can improve ease of use.

- It is necessary to develop a Neutral Language Processing Algorithm that will understand what the user is writing in the search box.

- It is necessary to design a 3D model where the user can choose where they see the symptoms.

4.3.1.1.2 The results of research 3 from the RiD perspective

A total of 219 comments were made. They were categorized into five. The thematic content is the comments that include the inferences the participants made using their knowledge of the given case story. As the name suggests, the medical content is all content that makes the system an expert. The UI/UX includes all comments made for interfaces, flows, user experience, and usability. The technological issues are the comments made by the participants about the problems they encounter due to the lack of some technologies in the app. The general comments are those that cannot be classified under other categories. The distribution of comments based on five categories made by each participant group is shown in Table 4.3.

As can be expected, the dentists made the most medical comments, and the designers made the most UI/UX comments (Figure 4.21). Since the parents are target users, they only commented on the problems they encountered while using the product. The reason why the dentists and the designers made more comments than the parents may be related to both their fields of expertise and their empathy.

Table 4.3: The distribution of the comments.

Participants	Thematic Content	Medical Content	UI/UX	Technological Issues	General
Parent 1	6	8	6	3	0
Parent 2	0	7	6	1	1
Parent 3	0	6	7	0	0
Dentist 1	0	13	11	0	1
Dentist 2	0	12	8	0	0
Dentist 3	2	14	11	1	0
Designer 1	0	6	33	2	1
Designer 2	0	7	19	0	0
Designer 3	0	6	21	0	0
SUM	8	79	122	7	3
TOTAL			219		

Dentists have sparked debate on medical context issues that I had not noticed before. Other participants also touched upon similar issues. However, the only way to make sure whether these issues were due to the lack of dentistry knowledge of the participants or because they were designed to be problematic, was to consult the dentists. For example, the parents and the designers asked what systemic disease means. The dentists, on the other hand, questioned how accurate the disease classification made in the profile was, and made recommendations based on their clinical experience, as they thought that it would not be understood by non-experts. They have also made suggestions that will improve the functioning of the algorithm beyond the medical content. This ability was related to the expertise they had.

The parents and the dentists also commented on the design. While they had difficulties due to problems in interfaces, usability, and user experience, the designers evaluated the design of the product in much more detail as it was related to their field of expertise and made recommendations for these issues, just as dentists made recommendations on the medical content.

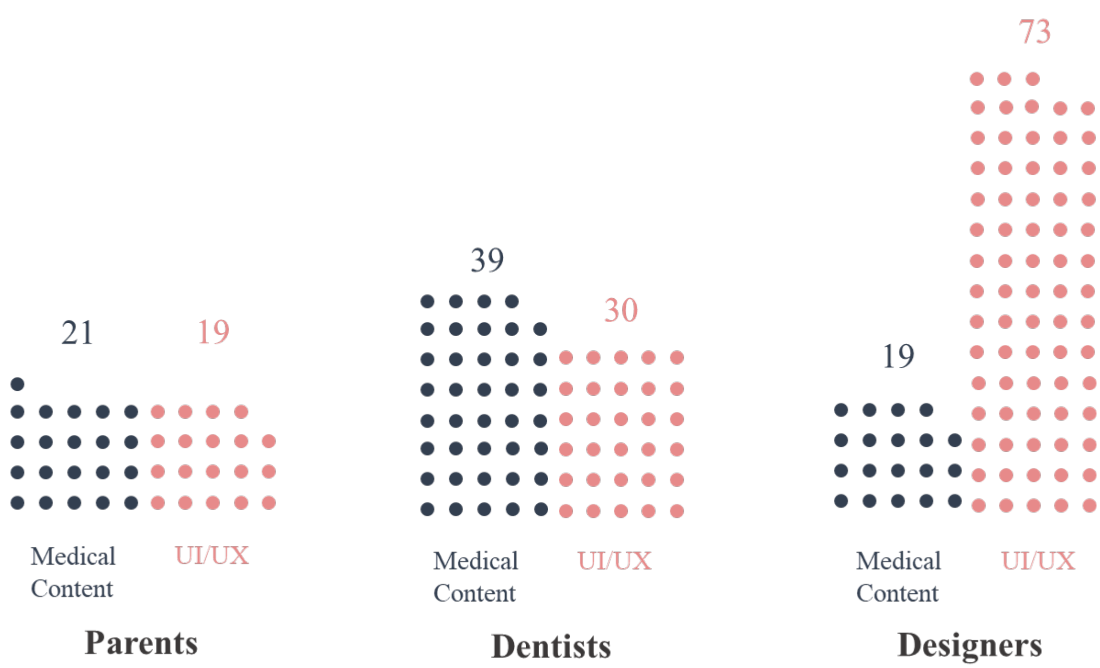


Figure 4.21: Number of medical content and UI/UX comments made by three participant groups.

This method was also useful to see the point of view of both sides or poles. For example, interesting comments about over-countered fillings were made by Parent 3 and Dentist 2 (the fillings that do not follow the form of tooth / misshaped fillings). After reading the question, "*Are there any over-countered fillings in your child's mouth that do not follow the form of the tooth?*", they said:

> Parent 3: Is it possible for a dentist to make such a filling?
>
> Dentist 2: In my experience, all fillings in children tend to overflow a bit.

I found these comments important in the sense that they reflect the perspectives of both parties regarding treatments.

Particularly surprising was the successful interpretation of Ali's data by Parent 1. At the beginning of the session, she said:

> "Ali is 7 years old, 40 kg, and 1.2 meters tall. He is quite heavy and tall for his age, like 12-13 years old. He has reflux as a systemic disease. He was drinking soda every day. That's why it seems normal to have reflux. He had a cleft palate and lip as a congenital disease. This child was most likely unable to breastfeed. He must also have low immunity because of it. He has allergies. It may have been because of his immunity. He likes sugary foods. Probably has caries too. He may have other addictions, possibly thumb-sucking. He was on an antihistamine. Is it something the doctor prescribed? Does he just take it only in springtime? That's what I thought."

Later, she answered the bedtime feeding question in the profile as follows:

> "Does he eat at bedtime? If he eats a lot of sugar, he gets hungry quickly. Probably he does."

During the assessment, she answered the question about vitamin deficiency as follows:

> "The child constantly consumes sugary foods; I think he has a vitamin deficiency."

These comments show that some parents, even if they are not experts, may have knowledge about the field and make inferences from their existing knowledge.

Considering all these results, it can be said that including dentists who are experts in medical content, designers who are experts in the field of UI/UX, and end users who are not experts in both fields, namely parents, enrich the data.

5. DISCUSSION

This thesis touches on different areas related to the project I have been carrying out for the last three years. Although the fields of dentistry, mobile applications, medical devices, expert systems, and human-computer interaction are extensive in themselves, since this thesis is in the design field, I tried to look at all of them from the design perspective. In addition to an extensive literature review in each field, I meticulously documented my design process, adopting a hybrid research approach. However, the thesis as a whole was strongly influenced by the principles of action research. Therefore, in the four subsections[17] of the literature review, my proposals, criticisms and comments on relevant subjects took place as a reflection. The most challenging part was writing such a complex study according to the school's current thesis template, where the author was expected to review the literature in the introduction and write only his/her statements in the discussion. In this thesis, you read some of my statements in the literature review. Moreover, although it is a design thesis, some sections that are not directly related to the design such as an introduction to expert systems, clinical assessment of oral health apps, or medical device regulations were included. In this sense, it is quite eclectic. Meyer and Batehup (1997) stress that being eclectic in action research is normal.

Why did I come up with this hybrid model? This was not what I originally planned. When I realized that there is no design study evaluating the usability of oral health mobile apps in the literature, I thought that I could do this research myself. However, at this time, I was in the research and development phase and had done design research for the algorithm, which was a unique study regarding the identification of oral diseases by

[17] (1) “Redefining the user in the context of expert systems”, (2) “A model proposition for profiling the users of expert systems”, (3) “The issues of approaches in the mDentistry literature” and (4) “Prospects and challenges of regulations”

the participants. When my advisor, Prof. Hümanur Bağlı, saw what I was doing, she encouraged me to write these studies as my thesis subject. But how would I handle this issue? Since I came from a dental education, I was not used to anything other than the positivist approach. On the other hand, it did not feel right to treat research as if I were an observer. On the contrary, I was the main actor in this process because I designed this product. I also thought that the research I did just to improve the product would not have significance in terms of literature. At the end of the day, I used the think-aloud method and there was nothing new in terms of design methodology. The only unusualness that existed in the research was closely related to the unusualness of the product. I was trying to develop an expert system for non-expert users. Even if I had determined my scope through HCI, I would not have presented a meaningful result in terms of literature because, as I said before, my main purpose in doing these studies was to develop the algorithm during the project. Moreover, it would be wrong to separate them from the entrepreneurship process. Baskerville (1999) says that complex systems cannot be simplified, even if they can be simplified, they would not be meaningful, therefore action research is the best way to study complex systems. At this point, I noticed that the principles of action research could provide me with a liberating space for documenting the complex process I was in.

While trying to figure out where to start, I came across the Ph.D. theses of two Danish researchers (Ludvigsen, 2006; Lykke-Olesen, 2006) who took action research as a valid stance and also benefited from the RtD approach. These researchers worked as designers in a new project that was being implemented in the Aarhus Library and wrote the whole process as their theses from different perspectives. These two theses became a reference point for me. The only difference is that after I completed the process, I looked back and reviewed my process. This was not hard as I also had the reports that I wrote for TÜBİTAK. Therefore, my approach can be described as semi-retrospective as well as hybrid. The reason why it is *semi*- is that the research was still going on when I started to write the thesis. As I reviewed the literature, new theories came to my mind, and thanks to this approach, I was able to document them. After I started writing, I also did the third research where the prototype in the form of an application was tested by the user which was also requested by the TÜBİTAK. In that sense, I killed two birds with one stone.

For the third study, Prof. Hümanur Bağlı suggested including users with three different backgrounds. Parents were the end users, but we thought the evaluations of dentists and designers would also enrich the data, and they did. When the method had to be evaluated as well, a purpose other than documenting an innovation process or developing the product emerged which made me realize the need for a hybrid approach.

When I read the article in which Frayling (1994) classified design research, I realized that research in these three classes has a place in my process. I considered the process as a whole and determined my position in the research as an actor or a participant. So, I was able to describe all my experiences and the decisions I made in an active voice. This fitted the definition of RtD. All three kinds of research I did in the process were RfD because their main objective was to develop the product. Involvement of the three pillars of the project in the third study: non-expert users, expert users, and designers was a new approach. Besides the design refinement purpose, I also intended to evaluate the effect of participants' expertise on their assessments, to see where they coincided and where they did not. This was obviously beyond the aim of the RfD approach and needed to be discussed with an outsider attitude. The evaluation has become product-independent, and it fit RiD. Therefore, I decided to bring all three together and propose a new hybrid design research approach. Prof. Hümanur Bağlı described this as the researcher's schizophrenia, but I've corrected this as a researcher's dissociative identity disorder, making it more appropriate to the medical term.

I also experienced an identity disorder in the entrepreneurial process itself. The TÜBİTAK project budget, which was quite limited, became even more limited due to the devaluation of the Turkish lira in November 2021 and the accompanying inflation. The budget was left to pay two people other than myself, and these two people were software developers. A software developer gave me a one-hour monthly consultation for a small fee. Other than that, I was alone in the R&D process. In addition to being a manager, medical engineer, UI designer, and design researcher in the project, I was the person who wrote the reports submitted to TUBITAK at regular intervals and applied for ethical approvals for the clinical study. In short, the multi-identity state in the entrepreneurship story continued during the research process.

It is stated that action research is a practical approach because it allows research in unpredictable conditions (Lewin, 1946; East & Robinson, 1994). When applying to the TUBITAK project, I had defined each stage, work packages and the actions to be taken. However, the stages in the R&D process were not "clean cut" as I stated in the project application (Figure 4.2). While developing the algorithm, I needed to do user research, and none of them was planned in advance and I had to start from the beginning. Similarly, I had planned to test the prototype in a clinical trial and after a while, I realized that the bureaucratic process required for permits would take a long time. The software process was delayed for a month because the software developers I interviewed did not want to sign a non-disclosure agreement. In short, no matter how pre-planned the process was, some conditions were constantly changing. Nothing was linear, but iterative (Figure 5.1). Meyer and Batehup (1997) argue that action research focuses on the process itself as well as the results of an innovation, thus allowing the change process to be explored rather than just aiming to measure output. Thanks to this principle, I was able to transfer the changes in my process.

What was the significance of the design for this project? As I mentioned earlier, expert systems are actually designed for the use of experts. If I were to design an expert system for them, it would be as important as the design is in any product. However, I was designing the product for non-expert users and since there was no such application available in the field of oral health, I had a tiger by the tail. How would I develop artificial intelligence to diagnose oral diseases? More importantly, how would I make it available to non-expert users? At this point, design played a key role.

In the process, I realized that it was not possible to design the system model exactly according to the user model. Considering that there were two poles on the axis of knowledge, on the one side, there was a non-expert user who knew nothing about the field, and on the other side an expert system in which knowledge about the field was concentrated. Parameters of the expert system, for example, as the definitions of symptoms approached the non-expert user, they lost their distinguishing features from other symptoms and started to resemble each other.

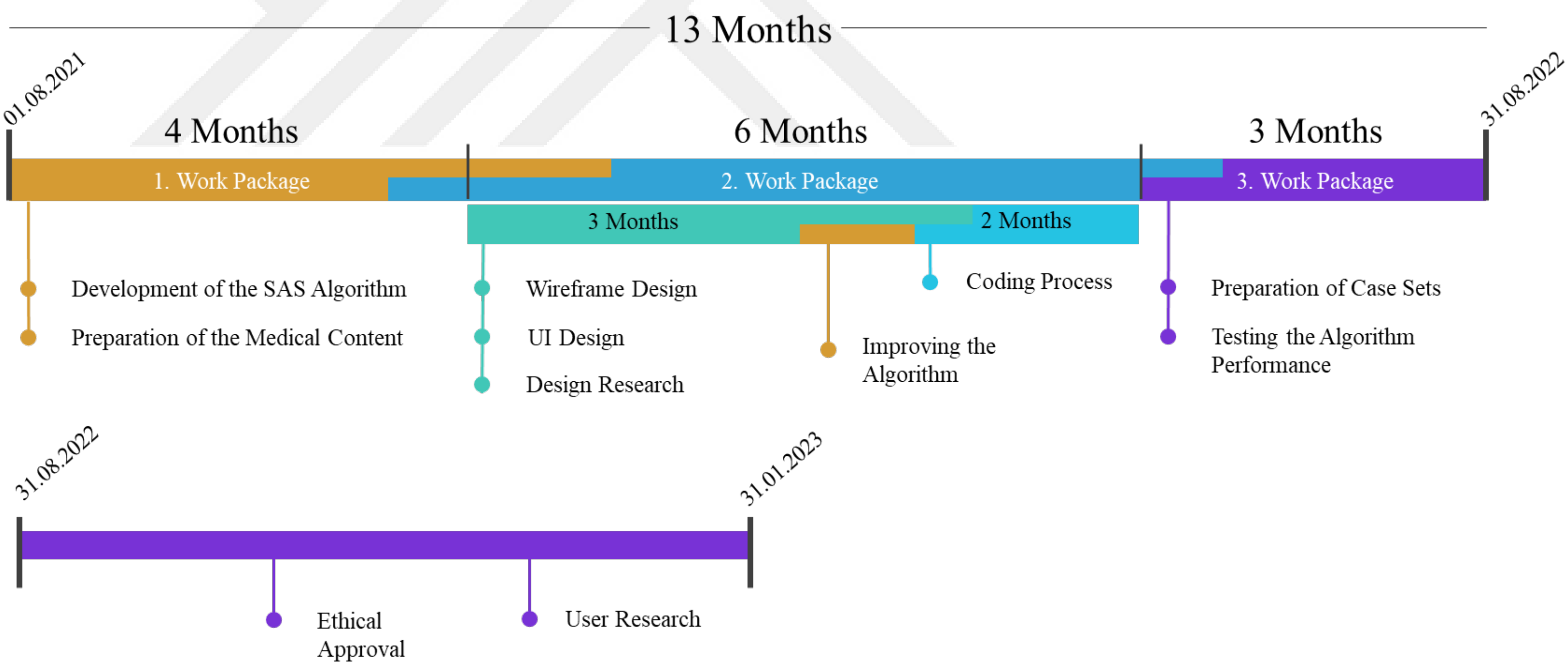


Figure 5.1: The actual process.

Therefore, there was a limit to this approach. In fact, the limit I defined for non-expert users was also valid for expert systems. Non-expert users had to make an effort to use an expert system and be willing to learn basic dental knowledge. In other words, the non-expert user needed to approach the expert system as much as the expert system approached the non-expert user. After identifying this situation, I started to consider whether my design process was user-centered, although users were included since the beginning. Perhaps, "User-approaching Design" better represents my design approach.

In her article (2010), Erin Friess examines how designers evaluate user research results and questions how much the user's wishes affect the design process. Criticizing that designers do not care if a user requests pink text, she states that design projects asserting that they are user-centered are not as user-centered as they claim. When I look at it from this perspective, I realize that I did not take the user at the center of my design process. If I asked them what they expected from an AI diagnosing oral diseases, they would probably be talking about an app that would tell them what happened when they held the phone's camera (I can say that I often encounter this misunderstanding when I tell people about my project). On the one hand, of course, I find this request reasonable, and I believe that the future goes in that direction. However, I should also point out that respondents do not always ask for reasonable things like the one who wants pink text. I can support this with the example of a designer participant in the second study where I tested symptom lists. He found the current list confusing and suggested that the symptoms should go hand in hand with their meanings. For example, according to him, all symptoms related to flow such as nasal discharge, pus discharge, and salivary discharge should be gathered under one title, but he was unaware of how much effort I had taken to collect 129 symptoms under 9 headings. If I did what he said, there would probably be dozens of titles, and this would be inaccurate from a medical perspective and make the algorithm more difficult to use. Therefore, I did not take his suggestion seriously (he would probably agree with me if he had worked in the research and development process). The contributions of the participants who do not have knowledge and experience in fields such as medicine or artificial intelligence passed through my filter. On this point, I agree with what Gaver (1991, p. 79) said: "We must understand the needs and abilities of prospective users. But equally, we must understand the capabilities and limitations of technologies to know the possibilities they offer for design." In the process of technological innovation, we should consider the needs of the user, but we should not overlook the fact that there are technical constraints and that the product is developed based on them in the context of a project.

As someone who entered the world of design from a different field, I liken this tension between engineers and designers to a tug-of-war game. Both sides have their own merits. While designers want users to be centered, engineers struggle with the technical challenges I mentioned above. I can predict the severity of this conflict, especially in R&D projects where the budget is limited. For my part, design education helped me see that the reductionist approach in natural sciences does not work very well in human-related issues. Of course, this approach is necessary to fix the parameters and produce scientific knowledge through this stability. However, as Susman and Evered (1978) say, when examining human organizations, this reductionist approach causes many things to be overlooked and the information produced by this approach cannot provide solutions to problems in practice. The person who encounters the product becomes a user and when we start talking about the user, many factors such as economic, socio-cultural, educational, etc. are involved in the process. In this context, it should be taken into account that a perspective considering a human being as a subject cannot produce valid information under the conditions in which these factors play a role. At such times, I remember what Prof. Sebnem Timur said in her "Cultural Approaches to Design" class: "Because you can't put culture in a test tube and examine it." For this reason, I find the efforts of the designers quite legitimate.

In fact, this situation also points to a problem that I often encounter in the entrepreneurship community, particularly in projects usually managed by engineers. During a conversation with an entrepreneur-engineer friend who had developed a product and was considering launching it, I discovered that she had never worked with a designer and had not yet tested the product with a target user. At this point, I find the studies conducted with medical device manufacturers (Privitera, 2017; Köse & Cifter, 2018) highly valuable in terms of due diligence. Although there are regulations (ISO

9241-210:2019) that recommend designing medical devices with a “human-centered” approach, the fact that manufacturers perceive designers as the staff that makes up the product, or at best work on human factors, shows us how important design and design research is not well-understood in the industry. However, assuming I am one of these manufacturers, I can easily say that design can do much more than that.

There was a critical moment in my design process while developing the algorithm. I created an algorithm from around 600 parameters that I extracted from hundreds of articles and organized them in a way I assumed the user would understand. At some point, I realized that I needed to test if these assumptions were true. So, I did the first research where I showed the participants some case photos and asked them to describe the symptoms they were seeing. Based on the results of this study, I went back to the beginning and rearranged some of the symptoms and the way the algorithm works. Therefore, the design research went beyond the interface and the human factor and was instrumental in changing the working principle of the inference engine. Based on this, I prepared a domination map.

As you can see on the map (Figure 5.2), in the expert system I developed, the dentistry was effective on the knowledge base, inference engine, and explanation facility, engineering was effective on the database, inference engine, and knowledge base, while the design influences user interface, explanation facility, and inference engine. I should point out here that the explanation facility does not function as a component that allows the developer to backtrack, as stated in the computer science literature (Gupta & Nagpal 2020, p.86). I defined this component as one that includes questions and explanations to extract more accurate data from the user which is more UX writing-related concern of the design.

It has been said that for the knowledge produced by action research to be valid, the research must be “*grounded*” (Susman & Evered, 1978). In other words, action research is not done to confirm or test a theory, but to explain or explore the existing situation and does not aim to produce scientific data. Action research is situation-specific and does not allow generalizations like scientific research. Some authors (Susman & Evered, 1978; Meyer & Batehup, 1997) stated that to close the gap between theory and practice, action researchers should generate a theory with the data they collect from practice. By doing so, it is emphasized that they should explain in detail what data they use, how, and why they use it should be transparent (Meyer and Batehup, 1997; Simons, 1981). In my thesis, I explained in detail the process of the innovation project in which I developed artificial intelligence in the field of oral health. As I conveyed the data I collected from the field, I clearly stated my reflection and the reasons for the decisions I made. I summarized the research results as much as possible to increase readability, but I tried to reflect the views of the participants in detail in order not to lose transparency. In this thesis, I never aimed to generate scientific data. That was not possible anyway. A unique process such as my innovation project could not be reduced to simple parts, generalizations could not be made, and the condition could not be reproduced by other researchers. For this reason, as Susman and Evered (1978) emphasized, my aim was to explore the process as a whole with the changes experienced. During the thesis process, I created theories by blending the literature I read from many fields with the experience and data I gained from the fieldwork. In this sense, I can say that I fulfilled some of the requirements of action research in my thesis.

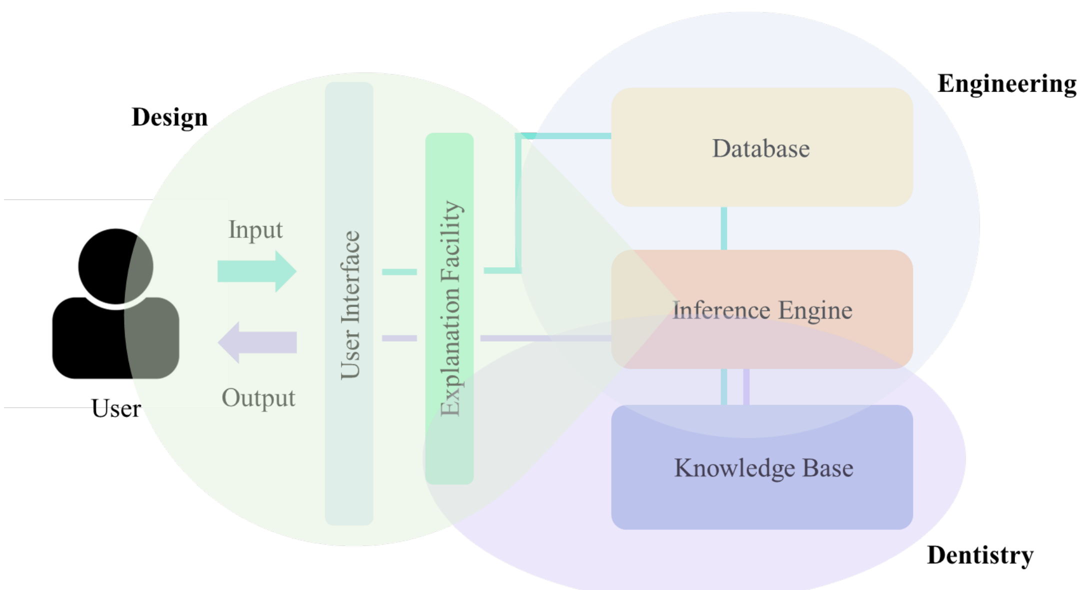


Figure 5.2: The map of dominance in the expert system I designed.

Susman and Evered (1978) describe cyclic processes that include five phases: 1) diagnosing, 2) action planning, 3) action taking, 4) evaluating and 5) specifying learning. The researcher observes the setting, identifies the problem, plans an intervention, intervenes, evaluates the results, and learns from his/her experience. These five phases are performed as many as needed until the researcher achieves the desired solution to the certain problem. It has a transformative dynamic to induce a change in human organizations by collaborating with the actors in the setting. Therefore, they claim that it is "*future-oriented*" and "*collaborative*." Considering these, what I did was not action research.

Nevertheless, I agree with Järvinen (2007) and Cole et al. (2005) about the fundamental similarities between design research and action research. Therefore, some of its principles fit my hybrid approach perfectly. Meyer (2000) points out that approaches can be shaped according to the unique situation of each researcher. For this reason, I would like to make it clear that I have applied these principles to my research style by being pragmatic, even though they do not fully comply with the forms in the literature.

5. CONCLUSIVE REMARKS

This study puts the design process of an expert system for non-expert users in the oral health field. in the center. By sprouting from there, it branches off to the HCI, the mDentistry, and the medical device field with the help of a comprehensive literature review.

It does not aim to produce scientific knowledge but to document a complex and unique process full of unpredictable situations and conflicting agendas.

For these purposes, a new hybrid model was created by combining each of Frayling's design research classes. In this way, the research and subsequent progress were able to be conveyed in harmony without being detached from its context and where I stand as a researcher.

In this thesis, some systematizations were developed and proposed such as the redefinition and reclassification of the users[18] and a model for user profiling in the context of expert systems or systems with expertise[19]; some criticisms and suggestions were made in terms of the adopted methodologies in the mDentistry literature[20]; and some prospects and challenges of regulatory enforcements[21] were addressed. The experiences gained from the innovation process and the comprehensive literature review played a role in their emergence and fed each of them to a great extent.

---

[18] Look at the sub section 2.1.1.2 Redefining the user in the context of expert systems.

[19] Look at the sub section 2.1.3.5 A model proposition for profiling the users of expert systems.

[20] Look at the sub section 2.2.1.5 The issues of approaches in the mDentistry literature.

[21] Look at the sub section 2.2.2.3 Prospects and challenges of regulations.

The R&D process of the project has not been completed yet. In the future, the findings of the third design research will be a great help in improving the algorithm, the interfaces, and the user experience.

Although this research does not aim to transform any problem setting (which is why it cannot be described as action research as defined in social studies), it has led to a transformation in me and a desire to rouse the actors from every field that this thesis touches. Therefore, I will begin by describing my personal learning experience as a brief self-reflection and finish off by calling out as a dentist, a medical engineer, an innovator, an entrepreneur, a medical device manufacturer, and a researcher positioned in two fields: healthcare and design.

My personal learning experience:

> As well as design education, this thesis changed my mind about the ways of generating knowledge and shook my belief that we can only reach the absolute truth through positivism. Because when I decided to write my innovation process as a thesis, I realized that complex processes with ever-changing conditions cannot be studied by being reduced to simple parts. In this world, where we continuously try to build knowledge structures, there is a vast of remnants that cannot be easily incorporated within those structures because they do not fit into certain patterns, they are intertwined. I believe they are still worth knowing. Approaches such as action research, RtD, or any type of practice-based research allow us to document them in their own unique manner. It emancipated me as a researcher. Without this approach, I could neither have criticized current classifications and methodologies, nor proposed new models.

To all researchers in the health field:

> I think that researchers in the health field should take advantage of the liberating opportunities of social studies more often. Of course, the knowledge we produce about the diagnosis or treatment methods of diseases must be obtained by scientific methods. The algorithm I have developed has emerged thanks to

evidence-based dentistry. However, it should not be overlooked that patients are human, and socio-cultural factors also play a role in healthcare. The way to better understand patients is through qualitative studies, not through the quantitative methods and statistical significance we are accustomed to in the field of medicine.

To the designers/engineers who design systems with expertise in all fields, especially healthcare, and to researchers who examine existing ones:

In this thesis, since I find the existing user classifications insufficient in the HCI and medical device literature, I reclassify the expert system users. Instead of the terms professional and lay-user, I divide users into experts and non-experts by taking the expertise of the product as a reference point. In this classification, unlike the existing ones, I categorize two main user types according to three axes: knowledge, experience, and competence. I think this may help us take a broader perspective when identifying users.

Nevertheless, I would like to remind you that I did not make this classification based on empirical data. I propose it as someone who has designed an expert system for non-expert users and reviewed the HCI and medical device literature. Another proposal I made is the model for creating user profiles in the context of an expert system or a system with expertise, which is in close relation with the user classification. This model can be leveraged to determine parameters related to participants while conducting usability research. Like the classification, this model was also not developed based on empirical data but on the experience and the knowledge I gained during the process and the comprehensive literature review.

However, it will only be possible to show whether the classification and the profiling model are valid in practice or other design processes, through further studies.

To policymakers working in the field of medical devices:

The fact that some mobile health applications are considered medical devices will increase the quality, safety, and traceability of the products in the market. In connection with this, the recent regulations expect manufacturers to prove that the performance and usability of their products are at a certain standard during the compliance and certification processes. It is stated in the literature that the studies on the clinical effectiveness of oral health apps lack methodological standards. In this sense, there is a need for clinical guidelines that will guide manufacturers in terms of proving the effectiveness and performance of their products. In addition, it is necessary to accelerate and facilitate compliance processes and make them entrepreneur-friendly to prevent the monopoly of large-scale companies so that small-scale manufacturers can survive in the industry.

To medical device manufacturers:

The results of the user research I conducted during the R&D process affected the inference engine of the algorithm. This shows us the power of design research. The design can change the working principles of the device beyond making it more appealing or arranging its ergonomics. For this reason, I argue that designers should be involved in the team from the very beginning of the development process.

To policymakers and the design academy:

Although we often see the phrase "human-centered design" in regulations, we understand how the role of a designer is perceived by medical device manufacturers and positioned in the industry thanks to the relevant studies. Whereas user-centered design is a delicate term that should be used more carefully. To reduce this gap between theory and practice, I coin the term "User-embracing Design" considering that we, manufacturers or innovators, may not be able to put all the wishes of the users at the center of the design in a challenging area dominated by the technology, health expertise and regulations. This term may be more congruent with the reality.

To health researchers working on oral health mobile applications and design researchers:

The studies on the clinical effectiveness of oral health mobile applications are not mature yet. There is a need to bring certain standards in terms of methodology to be able to make comparisons and meta-analyses. What if we were unable to demonstrate their effectiveness even if consensus was reached on these clinical research standards? What if the positivist approach were incapable of showing those apps' utility? Since laboratory conditions do not match real-life conditions, we may not be able to get meaningful results. Would this mean that oral health apps were useless? We need a different perspective here.

Moreover, the parameters in the proposed profiling model come into play when the participants become users rather than patients. The user characteristics affect the way of using the applications and thus their clinical performance. The current clinical studies ignore such differences between users. Without considering use and usability, one pillar will always be missing.

A kind reminder: Usability studies on oral health applications cannot be done without including the end users. Quantitative research methods are also insufficient for this purpose. Whereas usability studies in the design field are mostly carried out using qualitative and in-depth methods such as the cultural probe, think-aloud, observation, interview, etc. When evaluating the usability of products, healthcare professionals need to adopt a designer's perspective beyond using design heuristics.

It is unimaginable to think of developing a concept for placebo apps without designers. Therefore, health experts who have intentions to design one, you will need them. It is their expertise.

For all these reasons, I believe that researchers from both fields must cooperate. This is my call to action.

REFERENCES

Abdulraheem, S., and Bondemark, L. (2018). Hawthorne effect reporting in orthodontic randomized controlled trials: truth or myth? Blessing or curse? *European journal of orthodontics*, *40*(5), 475–479. https://doi.org/10.1093/ejo/cjx089

Abu-Nasser, B. S. (2017). Medical Expert Systems Survey. *International Journal of Engineering*, *1*(7), 218–224. Retrieved from http://ijeais.org/wp-content/uploads/2017/09/IJEAIS170918.pdf

ACM SIGCHI (1992). *Curricula for Human-computer Interaction*. ACM Special Interest Group on Computer-Human Interaction Curriculum Development Group, New York.

Agarwal, P., Mukerji, G., Desveaux, L., Ivers, N. M., Bhattacharyya, O., Hensel, J. M., Shaw, J., Bouck, Z., Jamieson, T., Onabajo, N., Cooper, M., Marani, H., Jeffs, L., & Bhatia, R. S. (2019). Mobile App for Improved Self-Management of Type 2 Diabetes: Multicenter Pragmatic Randomized Controlled Trial. *JMIR mHealth and uHealth*, *7*(1), e10321. https://doi.org/10.2196/10321

Aggarwal, R., Sounderajah, V., Martin, G., Ting, D. S. W., Karthikesalingam, A., King, D., Ashrafian, H., & Darzi, A. (2021). Diagnostic accuracy of deep learning in medical imaging: a systematic review and meta-analysis. *NPJ digital medicine*, *4*(1), 65. https://doi.org/10.1038/s41746-021-00438-z

Ahmed, N., Abbasi, M. S., Zuberi, F., Qamar, W., Halim, M. S. B., Maqsood, A., & Alam, M. K. (2021). Artificial Intelligence Techniques: Analysis, Application, and Outcome in Dentistry—A Systematic Review. *BioMed Research International*, *2021*, 1–15. https://doi.org/10.1155/2021/9751564

Akmal Muhamat, N., Hasan, R., Saddki, N., Mohd Arshad, M. R., & Ahmad, M. (2021). Development and usability testing of mobile application on diet and oral health. *PLOS ONE*, *16*(9), e0257035. https://doi.org/10.1371/journal.pone.0257035

Alkadhi, O. H., Zahid, M. N., Almanea, R. S., Althaqeb, H. K., Alharbi, T. H., & Ajwa, N. M. (2017). The effect of using mobile applications for improving oral hygiene in patients with orthodontic fixed appliances: a randomised controlled trial. *Journal of Orthodontics*, *44*(3), 157–163. https://doi.org/10.1080/14653125.2017.1346746

Alkilzy, M., Midani, R., Höfer, M., & Splieth, C. (2019). Improving Toothbrushing with a Smartphone App: Results of a Randomized Controlled Trial. *Caries Research*, *53*(6), 628–635. https://doi.org/10.1159/000499868

AlKlayb, S. A., Assery, M. K., AlQahtani, A., AlAnazi, M., & Pani, S. C. (2017). Comparison of the Effectiveness of a Mobile Phone-based Education Program in Educating Mothers as Oral Health Providers in Two Regions of Saudi Arabia. *Journal of International Society of Preventive & Community Dentistry*, *7*(3), 110–115. https://doi.org/10.4103/jispcd.JISPCD_95_17

Allwood, C. M. (1986). Novices on the computer: a review of the literature. *International Journal of Man-Machine Studies*, *25*(6), 633–658. https://doi.org/10.1016/s0020-7373(86)80079-7

Al-Musawi, A., Al-Sane, M., & Andersson, L. (2016). Smartphone App as an aid in the emergency management of avulsed teeth. *Dental Traumatology*, *33*(1), 13–18. https://doi.org/10.1111/edt.12298

Alqarni, A., Alfaifi, H., Aseeri, N., Gadah, T., & Togoo, R. (2018). Efficacy of a self-designed mobile application to improve child dental health knowledge among parents. *Journal of International Society of Preventive and Community Dentistry*, *8*(5), 424. https://doi.org/10.4103/jispcd.jispcd_195_18

American Academy of Pediatric Dentistry (2020a). Caries-risk assessment and management for infants, children, and adolescents. The Reference Manual of Pediatric Dentistry. Chicago, Ill.: *American Academy of Pediatric Dentistry*; 243-7.

American Academy of Pediatric Dentistry (2020b). Policy on social determinants of children's oral health and health disparities. The Reference Manual of Pediatric Dentistry. Chicago, Ill.: *American Academy of Pediatric Dentistry*; 28-31.

Anandakrishna, L., & Chandra, P. (2012). Is there an Association between Oral Health Status and School Performance? A Preliminary Study. *International Journal of Clinical Pediatric Dentistry*, *5(2)*, 132–135. https://doi.org/10.5005/jp-journals-10005-1150

Andersen, P. B. (1997). A theory of computer semiotics: *Semiotic approaches to construction and assessment of computer systems* (2nd ed.). Cambridge: Cambridge University Press.

Andersen, P. B. and Nowack P. (2002). Tangible Objects: Connecting Physical and Informational Space. *In Virtual space: spatiality in virtual inhabited 3D worlds*. 190 – 210. London, Springer-Verlag.

Andersson, L., Andreasen, J. O., Day, P., Heithersay, G., Trope, M., Diangelis, A. J., Kenny, D. J., Sigurdsson, A., Bourguignon, C., Flores, M. T., Hicks, M. L., Lenzi, A. R., Malmgren, B., Moule, A. J., Tsukiboshi, M., & International Association of Dental Traumatology (2012). International Association of Dental Traumatology guidelines for the management of traumatic dental injuries: 2. Avulsion of permanent teeth. *Dental traumatology: Official publication of International Association for Dental Traumatology*, *28*(2), 88–96. https://doi.org/10.1111/j.1600-9657.2012.01125.x

Andreasen, J. O., Bakland, L. K., & Andreasen, F. M. (2006). Traumatic intrusion of permanent teeth. Part 3. A clinical study of the effect of treatment variables such as treatment delay, method of repositioning, type of splint, length of splinting and antibiotics on 140 teeth. *Dental Traumatology*, *22*(2), 99–111. https://doi.org/10.1111/j.1600-9657.2006.00423.x

Antunes, J. L., Peres, M. A., de Campos Mello, T. R., & Waldman, E. A. (2006). Multilevel assessment of determinants of dental caries experience in Brazil. *Community dentistry and oral epidemiology*, *34*(2), 146–152. https://doi.org/10.1111/j.1600-0528.2006.00274.x

Apple Computer Inc. (1992). Macintosh Human Interface Guidelines. Addison-Wesley Professional.

Archer, B (1995). The nature of research. In Grand, S., Jonas, W (eds.) (2012) (eds.) *Mapping Design Research*. Birkhauser, Basel.

Ay, Z. Y., Sayin, M. O., Ozat, Y., Goster, T., Atilla, A. O., & Bozkurt, F. Y. (2007). Appropriate oral hygiene motivation method for patients with fixed appliances. *The Angle orthodontist*, *77*(6), 1085–1089. https://doi.org/10.2319/101806-428.1

Azarpazhooh, A., & Leake, J. L. (2006). Systematic review of the association between respiratory diseases and oral health. *Journal of periodontology, 77*(9), 1465–1482. https://doi.org/10.1902/jop.2006.060010

Barbosa, S. D. J., de Souza, C. S. and Prates, R. O. (2001). A Semiotic Engineering Approach to HCI. *Proc. Conference on Human Factors in Computing Systems CHI '01 extended abstracts on Human factors in computing systems*. 55 – 56. Seattle, Washington. ACM Press New York, NY.

Barnett, I., Torous, J., Staples, P., Sandoval, L., Keshavan, M., & Onnela, J. P. (2018). Relapse prediction in schizophrenia through digital phenotyping: a pilot study. *Neuropsychopharmacology: official publication of the American College of Neuropsychopharmacology*, *43*(8), 1660–1666. https://doi.org/10.1038/s41386-018-0030-z

Barr, P., Noble, J. and Biddle, R. (2002). *Icons R Icons: User interface icons, metaphor, and metonymy*. Technical Report CS-TR-02/20. Victoria University of Wellington, New Zealand. Retrieved from http://www.mcs.vuw.ac.nz/comp/Publications/archive/CS-TR-02/CS-TR-02-24.pdf

Barret, E., and Lally, V. (1999). Gender differences in an online learning environment. *Journal of Computer Assisted Learning, 15*(1), 48-60.

Baskerville, R. L. (1999). Investigating Information Systems with Action Research. *Communications of the Association for Information Systems*, 2. https://doi.org/10.17705/1cais.00219

Beck, J. D., Elter, J. R., Heiss, G., Couper, D., Mauriello, S. M., & Offenbacher, S. (2001). Relationship of periodontal disease to carotid artery Intima-Media wall thickness. *Arteriosclerosis, Thrombosis, and Vascular Biology, 21*(11), 1816–1822. https://doi.org/10.1161/hq1101.097803

Beck, J. D., Eke, P., Lin, D., Madianos, P., Couper, D., Moss, K., Elter, J., Heiss, G., & Offenbacher, S. (2005). Associations between IgG antibody to oral organisms and carotid intima-medial thickness in community-dwelling adults. *Atherosclerosis*, *183*(2), 342–348. https://doi.org/10.1016/j.atherosclerosis.2005.03.017

Begum, S., Ahmed, M. U., Funk, P., Xiong, N., & Folke, M. (2011). Case-Based Reasoning Systems in the Health Sciences: A Survey of Recent Trends and Developments. *IEEE Transactions on Systems, Man and Cybernetics*, *41*(4), 421–434. https://doi.org/10.1109/tsmcc.2010.2071862

Benyon, D., & Murray, D. (1993). Applying user modeling to human-computer interaction design. *Artificial Intelligence Review, 7*(3–4), 199–225. https://doi.org/10.1007/bf00849555

Bernabe, E., Marcenes, W., Hernandez, C. R., Bailey, J., Abreu, L. G., Alipour, V., Amini, S., Arabloo, J., Arefi, Z., Arora, A., Ayanore, M. A., Bärnighausen, T. W., Bijani, A., Cho, D. Y., Chu, D. T., Crowe, C. S., Demoz, G. T., Demsie, D. G., Dibaji Forooshani, Z. S., … Kassebaum, N. J. (2020). Global, Regional, and National Levels and Trends in Burden of Oral Conditions from 1990 to 2017: A Systematic Analysis for the Global Burden of Disease 2017 Study. *Journal of dental research*, *99*(4), 362–373. https://doi.org/10.1177/0022034520908533

Berthelot, J. M., Nizard, J., & Maugars, Y. (2019). The negative Hawthorne effect: Explaining pain overexpression. *Joint Bone Spine, 86*(4), 445–449. https://doi.org/10.1016/j.jbspin.2018.10.003

Birur, P. N., Sunny, S. P., Jena, S., Kandasarma, U., Raghavan, S., Ramaswamy, B., . . . Kuriakose, M. A. (2015). Mobile health application for remote oral cancer surveillance. *The Journal of the American Dental Association*, *146*(12), 886–894. https://doi.org/10.1016/j.adaj.2015.05.020

Biswas, P., & Springett, M. (2018). User modeling. (K. L. Norman & J. Kirakowski, Eds.), *The Wiley Handbook of Human-Computer Interaction* (1st ed., Vol. 1, pp. 143–170). John Wiley & Sons Ltd.

Biswas, P., Robinson, P. N., & Langdon, P. (2012). Designing Inclusive Interfaces Through User Modeling and Simulation. *International Journal of Human-computer Interaction, 28*(1), 1–33. https://doi.org/10.1080/10447318.2011.565718

Bolchini, D., Chatterji, R., & Speroni, M. (2009). Developing heuristics for the semiotics inspection of websites. *Proceedings of the 27th ACM*

International Conference on Design of Communication. https://doi.org/10.1145/1621995.1622009

Botwinick, J. (1973). *Aging and behaviour*. Springer, New York.

Bovair, S., Kieras, D. E., & Polson, P. G. (1990). The acquisition and performance of text-editing skill: A cognitive complexity analysis. *Human-Computer Interaction 5*, 1–48.

Bowen, B.T., Rinchuse, D. J., Zullo, T., & DeMaria, M. E. (2014). The influence of text messaging on oral hygiene effectiveness. *The Angle Orthodontist, 85*(4), 543–548. https://doi.org/10.2319/071514-495.1

Bray, F., Ferlay, J., Soerjomataram, I., Siegel, R. L., Torre, L. A., & Jemal, A. (2018). Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: A Cancer Journal for Clinicians*, *68*(6), 394–424. https://doi.org/10.3322/caac.21492

Brinker, T. J., Hekler, A., Enk, A. H., Klode, J., Hauschild, A., Berking, C., et al. (2019a). A Convolutional Neural Network Trained with Dermoscopic Images Performed on Par with 145 Dermatologists in a Clinical Melanoma Image Classification Task. *Eur. J. Cancer 111*, 148–154. doi:10.1016/j.ejca.2019.02.005

Brinker, T. J., Hekler, A., Enk, A. H., Klode, J., Hauschild, A., Berking, C., et al. (2019b). Deep Learning Outperformed 136 of 157 Dermatologists in a Head-To-Head Dermoscopic Melanoma Image Classification Task. *Eur. J. Cancer 113*, 47–54. doi:10.1016/j.ejca.2019.04.001

Broos, A., (2005). Gender and information and communication technologies (ICT) anxiety: Male self-assurance and female hesitation. *Cyber Psychology & Behaviour*, *8*(1), 21-31.

Burgess, C., & Swigger, K. (1986). A graphical database interface for casual, naive users. *Information Processing &Amp; Management, 22*(6), 511–521. https://doi.org/10.1016/0306-4573(86)90102-0

Busch, T. (1995). Gender differences in self-efficacy and attitudes toward computers. *Journal of Educational Computing Research 12*:147–158.

Campos, L. F., Cavalcante, J. P., Machado, D. P., Marçal, E., Silva, P. G. D. B., & Rolim, J. P. (2019). Development and Evaluation of a Mobile Oral Health Application for Preschoolers. *Telemedicine and e-Health, 25*(6), 492–498. https://doi.org/10.1089/tmj.2018.0034

Canbazoglu, E., Salman, Y. B., Yildirim, M. E., Merdenyan, B., & Ince, I. F. (2016). Developing a mobile application to better inform patients and enable effective consultation in implant dentistry. *Computational and Structural Biotechnology Journal, 14*, 252–261. https://doi.org/10.1016/j.csbj.2016.06.006

Card, S., Moran, T., & Newell, A. (1983). *The psychology of human-computer interaction*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Carrouel, F., Bourgeois, D., Clément, C., Tardivo, D., Martinon, P., Guiral, S., . . . Fraticelli, L. (2022). Oral-Hygiene-Related Mobile Apps in the French App Stores: Assessment of Functionality and Quality. International *Journal of Environmental Research and Public Health, 19*(12), 7293. https://doi.org/10.3390/ijerph19127293

Chapanis, A., (1959). *Research techniques in human engineering*. John Hopkins University Press, Baltimore, MA.

Chappuis, V., & von Arx, T. (2005). Replantation of 45 avulsed permanent teeth: a 1-year follow-up study. *Dental Traumatology, 21*(5), 289–296. https://doi.org/10.1111/j.1600-9657.2005.00330.x

Chen, M. (1986). Gender and computers: the beneficial effects of experience on attitudes. *Journal of Educational Computing Research 2*:265–282.

Chen, R., Santo, K., Wong, G., Sohn, W., Spallek, H., Chow, C., & Irving, M. (2021). Mobile Apps for Dental Caries Prevention: Systematic Search and Quality Evaluation. *JMIR mHealth and uHealth, 9*(1), e19958. https://doi.org/10.2196/19958

Chinn, C. H., Levine, J., Matos, S., Findley, S., & Edelstein, B. L. (2013). An Interprofessional Collaborative Approach in the Development of a Caries Risk Assessment mobile tablet application: My Smile Buddy. *Journal of Health Care for the Poor and Underserved, 24*(3), 1010–1020. https://doi.org/10.1353/hpu.2013.0114

Chou, J., & Hsiao, S. (2007). A usability study on human–computer interface for middle-aged learners. *Computers in Human Behavior, 23*(4), 2040–2063. https://doi.org/10.1016/j.chb.2006.02.011

Cianciola, L. J., Park, B. H., Bruck, E., Mosovich, L., & Genco, R. J. (1982). Prevalence of periodontal disease in insulin-dependent diabetes mellitus (juvenile diabetes). *Journal of the American Dental Association (1939), 104*(5), 653–660. https://doi.org/10.14219/jada.archive.1982.0240

Cifter, A. S., & Dong, H. (2009). Focusing on Users. *Include 09 Proceedings*. London: Helen Hamlyn Centre, Royal College of Art.

Cifter, A. S. (2011). An inclusive approach towards designing medical devices for use in the home environment. (Doctoral dissertation) Retrieved from https://bura.brunel.ac.uk/bitstream/2438/5515/1/FulltextThesis.pdf

Clemmensen, T. and Roese, K. (2010). An Overview of a Decade of Journal Publications about Culture and Human-Computer Interaction (HCI). In: Katre, D., Orngreen, R., Yammiyavar, P., Clemmensen, T. (eds) Human Work Interaction Design: Usability in Social, Cultural and Organizational Contexts. HWID 2009. *IFIP Advances in Information and Communication Technology*, vol 316. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-11762-6_9

Clemmensen, T., Hertzum, M., Hornbæk, K., Shi, Q., & Yammiyavar, P. (2009). Cultural cognition in usability evaluation. *Interacting With Computers, 21*(3), 212–220. https://doi.org/10.1016/j.intcom.2009.05.003

Cole, R., Purao, S., Rossi, M., & Sein, M. (2005). Being Proactive: Where Action Research Meets Design Research. *International Conference on Interaction Sciences 2005 Proceedings 27.* Retrieved from https://aisel.aisnet.org/icis2005/27

Coombs, M. J., & Alty, J. L. (1980). Face-to-face guidance of university computer users—II. Characterizing advisory interactions. *International Journal of Man-machine Studies, 12*(4), 407–429. https://doi.org/10.1016/s0020-7373(80)80024-1

Corston, R., & Colman, A. M. (1996). Gender and social facilitation effects on computer competence and attitudes toward computers. *Journal of Educational Computing Research 14*:171–183. https://doi.org/10.2190/7VW3-W6RV-6DCP-70MN

Cugelman, B. (2013). Gamification: What It Is and Why It Matters to Digital Health Behavior Change Developers. *JMIR Serious Games, 1*(1), e3. https://doi.org/10.2196/games.3139

Czaja, S. J., Sharit, J., (1993). Age differences in the performance of computer-based work. *Psychology and Aging, 8*(1), 59-67. https://doi.org/10.1037/0882-7974.8.1.59

Daher, J., Vijh, R., Linthwaite, B., Dave, S., Kim, J., Dheda, K., … Pai, N.P. (2017). Do digital innovations for HIV and sexually transmitted infections work? Results from a systematic review (1996-2017). *British Medical Journal Open, 7*, e017604. https://doi.org/10.1136/bmjopen-2017-017604

Daniel, S. J., & Kumar, S. (2014). Teledentistry: A Key Component in Access to Care. *Journal of Evidence-Based Dental Practice, 14*, 201–208. https://doi.org/10.1016/j.jebdp.2014.02.008

Darin, E. R., Meitzler, T. J., Witus, G., Sohn, E., Bryk, D., Goetz, R., and Gerhart, G. R., (2000). Computational modelling of age-differences in a visually demanding driving task: vehicle detection. *IEEE Transactions on Systems, Man, and Cybernetics, Part A: Systems and Humans*, Vol.30, No.3., 336-346.

Davenport, E. S., Williams, C. E., Sterne, J. A., Murad, S., Sivapathasundram, V., & Curtis, M. A. (2002). Maternal periodontal disease and preterm low birthweight: case-control study. *Journal of dental research, 81*(5), 313–318. https://doi.org/10.1177/154405910208100505

Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *Management Information Systems Quarterly, 13*(3), 319. https://doi.org/10.2307/249008

de Souza, C. S. (2005). *The semiotic engineering of human-computer interaction*. Cambridge, MA: MIT Press.

de Souza, C. S. (2018). Semiotics and Human-Computer Interaction. (K. L. Norman & J. Kirakowski, Eds.), *The Wiley Handbook of Human-Computer Interaction* (1st ed., Vol. 1, pp. 33–50). John Wiley & Sons Ltd.

Deleuse, M., Meiffren, C., Bruwier, A., Maes, N., Le Gall, M., & Charavet, C. (2020). Smartphone application-assisted oral hygiene of orthodontic patients: a multicentre randomized controlled trial in adolescents. *European Journal of Orthodontics, 42*(6), 605–611. https://doi.org/10.1093/ejo/cjz105

den Boer, J., van Dijk, W., Horn, V., Hescot, P., & Bruers, J. (2018). Collecting standardised oral health data via mobile application: A proof of concept study in the Netherlands. *PloS one, 13*(2), e0191385. https://doi.org/10.1371/journal.pone.0191385

Deshpande, S., Chahande, J., & Rathi, A. (2017). Mobile learning app: A novel method to teach clinical decision making in prosthodontics. *Education for Health, 30*(1), 31. https://doi.org/10.4103/1357-6283.210514

Desvarieux, M., Demmer, R. T., Rundek, T., Boden-Albala, B., Jacobs, D. R., Sacco, R. L., & Papapanou, P. N. (2005). Periodontal Microbiota and Carotid Intima-Media Thickness. *Circulation, 111*(5), 576–582. https://doi.org/10.1161/01.cir.0000154582.37101.15

Dong, J., & Salvendy, G. (1999). Designing menus for the Chinese population: Horizontal or vertical? *Behaviour & Information Technology, 18*(6), 467–471. https://doi.org/10.1080/014492999118887

Doyle, J. R. (1990). Naïve users and the Lotus* interface: a field study. *Behaviour &Amp; Information Technology, 9*(1), 81–89. https://doi.org/10.1080/01449299008924226

Dunn, J. L., Ko, K. K., Lahoud, D., Nusem, E., Straker, K., & Wrigley, C. (2019) Exploring the role of design in the context of medical device innovation. In Almendra, R.A. (Ed.) *Proceedings of the Academy for Design Innovation Management Conference 2019*. Academy for Design Innovation Management, United Kingdom, pp. 1-19.

Eason, K. (1974). The manager as a computer user. *Applied Ergonomics*. https://doi.org/10.1016/0003-6870(74)90252-x

East, L., & Robinson, J. (1994). Change in process: bringing about change in health care through action research. *Journal of Clinical Nursing, 3*(1), 57–61. https://doi.org/10.1111/j.1365-2702.1994.tb00359.x

Eco, U. (1976). *A Theory of Semiotics*, Indiana University Press, Bloomington, IN.

Eco, U. (1988). *Le signe: histoire et analyse d'un concept*. Retrieved May 2, 2023, from http://www.signosemio.com/eco/semiotic-process-and-classification-of-signs.asp

Edwards, E. A., Lumsden, J., Rivas, C., Steed, L., Edwards, L. A., Thiyagarajan, A., . . . Walton, R. T. (2016). Gamification for health promotion: systematic review of behaviour change techniques in smartphone apps.

BMJ Open, 6(10), e012447. https://doi.org/10.1136/bmjopen-2016-012447

Elliott, J. (1991). *Action research for educational change* (Theory in Practice) (1st ed.). Open University Press.

Engebretson, S. P., Lamster, I. B., Elkind, M. S., Rundek, T., Serman, N. J., Demmer, R. T., . . . Desvarieux, M. (2005). Radiographic Measures of Chronic Periodontitis and Carotid Artery Plaque. *Stroke, 36*(3), 561–566. https://doi.org/10.1161/01.str.0000155734.34652.6c

Engelbrektsson, P., Yesil, Ö, and Karlsson, I. C. M. (2000). Eliciting customer requirements in focus group interviews: can efficiency be increased? *EIASM: The 7th international product development management conference*, May 29-30. 129-141

Essén, A., Stern, A. D., Haase, C. B., Car, J., Greaves, F., Paparova, D., . . . Bates, D. W. (2022). Health app policy: international comparison of nine countries' approaches. *Npj Digital Medicine, 5*(1). https://doi.org/10.1038/s41746-022-00573-1

Estai, M., Kanagasingam, Y., Huang, B., Checker, H., Steele, L., Kruger, E., & Tennant, M. (2016). The efficacy of remote screening for dental caries by mid-level dental providers using a mobile teledentistry model. *Community Dentistry and Oral Epidemiology, 44*(5), 435–441. https://doi.org/10.1111/cdoe.12232

Estai, M., Kanagasingam, Y., Huang, B., Shiikha, J., Kruger, E., Bunt, S., & Tennant, M. (2017). Comparison of a Smartphone-Based Photographic Method with Face-to-Face Caries Assessment: A Mobile Teledentistry Model. *Telemedicine and e-Health, 23*(5), 435–440. https://doi.org/10.1089/tmj.2016.0122

European Commission. (n.d.). *Bodies, legislation: Regulation (EU) 2017/745 on medical devices*. Retrieved November 16, 2022, from https://ec.europa.eu/growth/tools-databases/nando/index.cfm?fuseaction=directive.notifiedbody&sort=country&dir_id=34

European Parliament. (2016). *OJ L 119 4.5.2016 Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation)*. Retrieved November 10, 2022, from https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:02016R0679-20160504

Fairbanks, R. J. & Wears, R. L. (2008). Hazards with medical devices: the role of design. *Annals of emergency medicine, 52*, pp.519–521.

Fairbanks, R. J., & Caplan, S. (2004). Poor interface design and lack of usability testing facilitate medical error. *Joint Commission journal on quality and safety*, *30*(10), 579–584. https://doi.org/10.1016/s1549-3741(04)30068-7

Farhadifard, H., Soheilifar, S., Farhadian, M., Kokabi, H., & Bakhshaei, A. (2020). Orthodontic patients' oral hygiene compliance by utilizing a smartphone application (Brush DJ): a randomized clinical trial. BDJ Open, 6(1). https://doi.org/10.1038/s41405-020-00050-5

Faulkner, L., and Wick, D. (2005). Cross-user analysis: Benefits of skill level comparison in usability testing. *Interacting with Computers, 17*, 773-786. https://doi.org/10.1016/j.intcom.2005.04.004

Fernández, C., Maturana, C., Coloma, S., Carrasco-Labra, A., & Giacaman, R. (2021). Teledentistry and mHealth for Promotion and Prevention of Oral Health: A Systematic Review and Meta-analysis. *Journal of Dental Research*, 002203452110038. https://doi.org/10.1177/00220345211003828

Ferreira, J., Barr, P., & Noble, J. (2005). The semiotics of user interface redesign. Australasian *User Interface Conference* (pp. 47–53). Retrieved from https://crpit.com/confpapers/CRPITV40Ferreira.pdf

Findeli, A. (2004). La recherche-projet: une méthode pour la recherche en design. *Paper presented at the Symposium de recherche sur le design*, Bâle, Suisse. Retrieved from https://projekt.unimes.fr/files/2014/04/Findeli.2005.Recherche-projet.pdf

Fischer, G. (2001). User modeling in Human-Computer Interaction. *User Modeling and User-adapted Interaction, 11*, 65–86. https://doi.org/10.1023/a:1011145532042

Fisher, J. & Kaplan, I. (1990). Representation of computer processes and command words, in L. Berlinguet and D. Berthelette (eds) *Work with display units 89* (Amsterdam: North Holland).

Fisher, J. (1991). Defining the novice user. *Behaviour &Amp; Information Technology, 10*(5), 437–441. https://doi.org/10.1080/01449299108924301

Ford, G., & Gelderblom, H. (2003). The effects of culture on performance achieved through the use of human computer interaction. *South African Institute of Computer Scientists and Information Technologists* (pp. 218–230). https://doi.org/10.5555/954014.954038

Fraser, M. (2013). *Design Research in Architecture: An Overview*, Surrey, Ashgate Publishing.

Frayling, C. (1994). Research in Art and Design. *Royal College of Art Research Papers*. London: Royal College of Art.

Friess, E. (2010). The Sword of Data: Does Human-Centered Design Fulfill Its Rhetorical Responsibility? *Design Issues, 26*(3), 40–50. https://doi.org/10.1162/desi_a_00028

Fung C.H., Igodan U., Cathy A., Martin J.L., Dzierzewski J.M., Josephson K. & Kramer J. (2015a). Human Factors/Usability Barriers to Home Medical Devices Among Individuals with Disabling Conditions: In-depth

Interviews with Positive Airway Pressure Device Users. *Disability and Health Journal, 8*, 86-92. https://doi.org/10.1016/j.dhjo.2014.06.002

Fung C.H., Martin J., Hays R.D., Rodriguez J.C., Igodan U., Jouldjian S., Dzierzewski J.M., Kramer B.J., Josephson K. & Alessi C. (2015b). Development of the Usability of Sleep Apnea Equipment – Positive Airway Pressure (USE-PAP) Questionnaire. *Sleep Medicine, 16*, 645-651. https://doi.org/10.1016/j.sleep.2015.01.019

Gaines, B. R. (1991). Designing Expert Systems for Usability. In B. Shackel & S. J. Richardson, (Eds.). *Human Factors for Informatics Usability* (pp. 207-247). Cambridge University Press.

Gaver, W.W. (1991). Technology affordances. *CHI'91 Conference Proceedings*. 79-84.

Gerritsen, A. E., Allen, P. F., Witter, D. J., Bronkhorst, E. M., & Creugers, N. H. (2010). Tooth loss and oral health-related quality of life: a systematic review and meta-analysis. *Health and Quality of Life Outcomes, 8*(1), 126. https://doi.org/10.1186/1477-7525-8-126

Gibbons, M., Limoges, C., Nowotny, H., Schwartzman, S., Scott, P., & Trow, M. (1994). *The New Production of Knowledge: The Dynamics of Science and Research in Contemporary Societies* (1st ed.). SAGE Publications Ltd.

Gibson, E. J. and Pick, A. D. (2003). *An Ecological Approach to Perceptual Learning and Development*. Cary, USA, Oxford University Press

Gibson, J. L. (1977). The theory of affordances. In Gieseking, J. J., Mangold, W., Katz, C., Low, S. M., & Saegert, S. (Ed.). *The People, Place, and Space Reader* (pp.56-60) Routledge.

Glanville, R. (1997). A Ship Without a Rudder. In: Glanville, R. & Zeeuw, G. D. (eds.) *Problems of Excavating Cybernetics and Systems*. Southsea: BKS+.

Goepfert, A. R., Jeffcoat, M. K., Andrews, W. W., Faye-Petersen, O., Cliver, S. P., Goldenberg, R. L., & Hauth, J. C. (2004). Periodontal disease and upper genital tract inflammation in early spontaneous preterm birth. *Obstetrics and gynecology, 104*(4), 777–783. https://doi.org/10.1097/01.AOG.0000139836.47777.6d

Goldberg, J., Privitera, M.B. & Malissagne, P. (2007). Benefits of Collaboration Between biomedical Engineering and Industrial Design Students in Senior Design Projects. *IEEE Engineering in Medicine and Biology Magazine*, May/June(4), pp.98–99.

Gomes, M. S., Bonan, P. R. F., Ferreira, V. Y. N., de Lucena Pereira, L., Correia, R. J. C., da Silva Teixeira, H. B., . . . Bonan, P. (2017). Development of a mobile application for oral cancer screening. *Technology and Health Care, 25*(2), 187–195. https://doi.org/10.3233/thc-161259

Goulding, A. (2001). Information Poverty or Overload? *Journal of Librarianship and Information Science, 33*(3), 109-111.

Graham, A. K., Greene, C. J., Kwasny, M. J., Kaiser, S. M., Lieponis, P., Powell, T., & Mohr, D. C. (2020). Coached Mobile App Platform for the

Treatment of Depression and Anxiety Among Primary Care Patients: A Randomized Clinical Trial. *JAMA psychiatry*, *77*(9), 906–914. https://doi.org/10.1001/jamapsychiatry.2020.1011

Greenwood, J. (1994). Action research: a few details, a caution and something new. *Journal of Advanced Nursing, 20*(1), 13–18. https://doi.org/10.1046/j.1365-2648.1994.20010013.x

Guillemette, L. and Cossette, J. (2006). *Mode of sign production* Retrieved May 2, 2023, from http://www.signosemio.com/eco/modes-of-sign-production.asp

Gupta, I., & Nagpal, G. (2020). *Artificial Intelligence and Expert Systems*. Mercury Learning and Information.

Hart, F., & Bond, M. (1995). *Action Research for Health and Social Care: A Guide to Practice*. Amsterdam, Netherlands: Amsterdam University Press.

Harte R. P., Glynn L. G., Broderick B. J., Rodriguez-Molinero A., Baker P. M. A., McGuiness B., O'Sullivan L., Diaz M., Quinlan L. R. & OLaighin G. (2014). Human-centered Design Considerations for Connected Health Devices for the Older Adult. *Journal of Personalized Medicine, 4*, 245-281. https://doi.org/10.3390/jpm4020245

Hatlevik, O. E., & Christophersen, K. (2013). Digital competence at the beginning of upper secondary school: Identifying factors explaining digital inclusion. *Computers & Education, 63*, 240–247. https://doi.org/10.1016/j.compedu.2012.11.015

Health on the Net. (2020). *The 8 principles of the HONcode app certification*. Retrieved October 30, 2022, from https://www.hon.ch/en/certification/app-certification-en.html

Ho, T. C. Y., McGrath, C., Yiu, C. K. Y., & Lee, G. H. M. (2022). Apps for Promoting Children's Oral Health: Systematic Search in App Stores and Quality Evaluation. *JMIR Pediatrics and Parenting, 5*(2), e28238. https://doi.org/10.2196/28238

Hockley, J., & Froggatt, K. (2006). The development of palliative care knowledge in care homes for older people: the place of action research. *Palliative Medicine, 20*(8), 835–843. https://doi.org/10.1177/0269216306073111

Hofstede, G. H. (1991). *Cultures and Organizations: Software of the Mind*. McGraw-Hill Companies.

Hogg, C., & Williamson, C. (2001). Whose interests do lay people represent? Towards an understanding of the role of lay people as members of committees. *Health expectations: an international journal of public participation in health care and health policy, 4*(1), 2–9. https://doi.org/10.1046/j.1369-6513.2001.00106.x

Holfelder, M., Mulansky, L., Schlee, W., Baumeister, H., Schobel, J., Greger, H., . . . Pryss, R. (2021). Medical Device Regulation Efforts for mHealth Apps during the COVID-19 Pandemic—An Experience Report of Corona

Check and Corona Health. *J*, *4*(2), 206–222. https://doi.org/10.3390/j4020017

Huckvale, K., Prieto, J. T., Tilney, M., Benghozi, P. J., & Car, J. (2015). Unaddressed privacy risks in accredited health and wellness apps: a cross-sectional systematic assessment. *BMC Medicine, 13*(1). https://doi.org/10.1186/s12916-015-0444-y

Hung, H. C., Joshipura, K. J., Colditz, G., Manson, J. E., Rimm, E. B., Speizer, F. E., & Willett, W. C. (2004). The association between tooth loss and coronary heart disease in men and women. *Journal of Public Health Dentistry, 64*(4), 209–215. https://doi.org/10.1111/j.1752-7325.2004.tb02755.x

Ihme, J. M., & Senkbeil, M. (2017). Warum können Jugendliche ihre eigenen computerbezogenen Kompetenzen nicht realistisch einschätzen? *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 49(1), 24–37. https://doi.org/10.1026/0049-8637/a000164

IMDRF SaMD Working Group (2013). *Software as a medical device (SaMD): Key definitions*. Retrieved November 8, 2022, from https://www.imdrf.org/documents/software-medical-device-samd-key-definitions

Islam, M. N. (2013). A systematic literature review of semiotics perception in user interfaces. *Journal of Systems and Information Technology, 15*(1), 45–77. https://doi.org/10.1108/13287261311322585

Islam, M. N. (2011). A semiotics perspective to web usability: an empirical case study. *Proceedings of the IADIS International Conference on Interfaces and Human Computer Interaction (IHCI 2011)*, Rome, Italy.

ISO (1998). *Ergonomic Requirements for Office Work with Visual Display Terminals (VDTs) – Guidance on Usability*, ISO 9241-11, International Organization for Standardization, Geneva.

ISO 9241-210:2019. (n.d.). Retrieved from https://www.iso.org/standard/77520.html

Jacobson, D., Jacobson, J., Leong, T., Lourenco, S., Mancl, L., & Chi, D. L. (2019). Evaluating Child Toothbrushing Behavior Changes Associated with a Mobile Game App: A Single Arm Pre/Post Pilot Study. *Pediatric dentistry, 41*(4), 299–303.

Jamieson, L., Elani, H., Mejia, G., Ju, X., Kawachi, I., Harper, S., . . . Kaufman, J. (2016). Inequalities in Indigenous Oral Health: Findings from Australia, New Zealand, and Canada. *Journal of Dental Research, 95*(12), 1375–1380. https://doi.org/10.1177/0022034516658233

Janhager, J. (2005). User consideration in early stages of product development – Theories and methods. Doctoral thesis. Royal Institute of Technology, Stockholm, Sweden.

Jarjoura, K., Devine, P. C., Perez-Delboy, A., Herrera-Abreu, M., D'Alton, M., & Papapanou, P. N. (2005). Markers of periodontal infection and preterm birth. *American journal of obstetrics and gynecology, 192*(2), 513–519. https://doi.org/10.1016/j.ajog.2004.07.018

Järvinen, P. (2007). Action Research is Similar to Design Science. *Quality & Quantity, 41*(1), 37–54. https://doi.org/10.1007/s11135-005-5427-1

Javaid, M., Ashrafi, S., Zefran, M., & Steinberg, A. D. (2016). ToothPIC: An Interactive Application for Teaching Oral Anatomy. *IEEE Transactions on Learning Technologies, 9*(2), 184–189. https://doi.org/10.1109/tlt.2015.2504963

John B. E., & Kieras, D. (1996). The GOMS family of user interface analysis techniques: Comparison and contrast. *ACM Transactions on Computer-Human Interaction, 3*, 320–351.

Johnson, E. M., and Baker, J. D. (1974). Field testing. The delicate compromise. *Human Factors, 16*, 203-214.

Jones, P. H. (2013). Design for Care: Innovating Healthcare Experience. Brooklyn, N.Y: Rosenfeld Media.

Jordan, P. W., Thomas, B., McClelland, I. L., Weerdmeester, B., & Brooke, J. (1996). SUS: a “quick and dirty” usability scale. *Usability evaluation in industry* (1st ed.). CRC Press.

Kadakia, K., Patel, B., & Shah, A. (2020). Advancing digital health: FDA innovation during COVID-19. *Npj Digital Medicine, 3*(1). https://doi.org/10.1038/s41746-020-00371-7

Kanoute, A., Carrouel, F., Gare, J., Dieng, S. N., Dieng, A., Diop, M., . . . Bourgeois, D. (2022). Evaluation of Oral Hygiene-Related Mobile Apps for Children in Sub-Saharan Africa. *International Journal of Environmental Research and Public Health, 19*(19), 12565. https://doi.org/10.3390/ijerph191912565

Kant, I. (1908) The Critique of Pure Reason (1781). in B. Rand, (ed.) *Modern Classical Philosophers*, Cambridge, MA: Houghton Mifflin, pp. 370-456.

Kaptelinin, V. (2014). *Affordances and Design*. (A. Campbell, Ed.) (1st ed.). The Interaction Design Foundation.

Kassebaum, N., Bernabé, E., Dahiya, M., Bhandari, B., Murray, C., & Marcenes, W. (2014). Global Burden of Severe Periodontitis in 1990-2010. *Journal of Dental Research, 93*(11), 1045–1053. https://doi.org/10.1177/0022034514552491

Kaye, R., & Crowley, J. (2000). *Medical Device Use-Safety: Incorporating Human Factors Engineering into Risk Management: Guidance for Industry and FDA Premarket and Design Control Reviews*. U.S. Food and Drugs Administration, Centre for Devices and Radiological Health.

Keutzer, L., & Simonsson, U. S. (2020). Medical Device Apps: An Introduction to Regulatory Affairs for Developers. *JMIR mHealth and uHealth, 8*(6), e17567. https://doi.org/10.2196/17567

Kientz, J. A., Choe, E. K., Birch, B., Maharaj, R., Fonville, A., Glasson, C., & Mundt, J. (2010). Heuristic evaluation of persuasive health technologies. *Proceedings of the ACM International Conference on Health Informatics - IHI '10*. https://doi.org/10.1145/1882992.1883084

Kieras, D. (1994). GOMS Modeling of user interfaces using NGOMSL. *Conference companion on human factors in computing systems* (pp. 371–372). Boston, MA: ACM.

Kieras, D. E., Wood, S. D., Abotel, K. & Hornof A. (1995). GLEAN: A computer-based tool for rapid GOMS model usability evaluation of user interface designs. *In Proceedings of ACM Symposium on User Interface and Software Technology* (pp. 91–100). New York, NY: ACM.

Kim, T. W., & Werbach, K. (2016). More than just a game: ethical issues in gamification. *Ethics and Information Technology, 18*(2), 157–173. https://doi.org/10.1007/s10676-016-9401-5

Kocaballi, A. B., Ijaz, K., Laranjo, L., Quiroz, J. C., Rezazadegan, D., Tong, H. L., . . . Coiera, E. (2020). Envisioning an artificial intelligence documentation assistant for future primary care consultations: A co-design study with general practitioners. *Journal of the American Medical Informatics Association, 27*(11), 1695–1704. https://doi.org/10.1093/jamia/ocaa131

Koch, G., Poulsen, S., Espelid, I., & Haubek, D. (2017). *Pediatric Dentistry: A Clinical Approach* (3rd ed.). Chichester, West Sussex, UK: Wiley-Blackwell.

Köse, N. G., & Cifter, A. S. (2018). Türkiye tıbbi cihazlar sektörüne yönelik bir inceleme: Ürün geliştirme süreçlerinin insan merkezli tasarım üzerinden değerlendirilmesi. *Online Journal of Art and Design, 6*:5, 239-262.

Kotecha, D., Chua, W. W. L., Fabritz, L., Hendriks, J., Casadei, B., Schotten, U., Vardas, P., Heidbuchel, H., Dean, V., Kirchhof, P., & European Society of Cardiology (ESC) Atrial Fibrillation Guidelines Taskforce, the CATCH ME consortium and the European Heart Rhythm Association (EHRA) (2018). European Society of Cardiology smartphone and tablet applications for patients with atrial fibrillation and their health care providers. *Europace : European pacing, arrhythmias, and cardiac electrophysiology : journal of the working groups on cardiac pacing, arrhythmias, and cardiac cellular electrophysiology of the European Society of Cardiology*, *20*(2), 225–233. https://doi.org/10.1093/europace/eux299

Kuniavsky, M. (2003). *Observing the User Experience: A Practitioner's Guide to User Research*. San Francisco: Morgan Kaufmann

Kurniawan, S. (2008). Older people and mobile phones: A multi-method investigation. *International Journal of Human-computer Studies, 66*(12), 889–901. https://doi.org/10.1016/j.ijhcs.2008.03.002

Kurzweil, R. (1992). *The Age of Intelligent Machines*. The MIT Press.

Laguna, K. D., & Babcock, R. L. (1997). Computer anxiety in young and older adults: Implications for human-computer interactions in older populations. *Computers in Human Behavior, 13*(3), 317–326. https://doi.org/10.1016/s0747-5632(97)00012-5

Landauer, T. K. (1995). *The Trouble with Computers: Usefulness, Usability, and Productivity*, Cambridge, MA: The MIT Press.

Larkin, J. H. (1981). Enriching formal knowledge: a model for learning to solve textbook physics problems. In Anderson, J. R. Ed., *Cognitive Skills and Their Acquisition*. Hillsdale, New Jersey: Erlbaum.

Larkin, J. H. (1983). The role of problem representation in physics. In Gentner, D. & Stevens, A. L. Eds, *Mental Models*. Hillsdale, New Jersey: Erlbaum.

Larson, J. L., Rosen, A. B., & Wilson, F. A. (2018). The Effect of Telehealth Interventions on Quality of Life of Cancer Patients: A Systematic Review and Meta-Analysis. *Telemedicine journal and e-health : the official journal of the American Telemedicine Association*, *24*(6), 397–405. https://doi.org/10.1089/tmj.2017.0112

Lee, K. H., Choi, Y. Y., & Jung, E. S. (2023). Effectiveness of an oral health education programme using a mobile application for older adults: A randomised clinical trial. *Gerodontology*, *40*(1), 47–55. https://doi.org/10.1111/ger.12616

Levin, T., & Gordon, C. (1989). Effect of gender and computer experience on attitudes toward computers. *Journal of Educational Computing Research* 5:69–88.

Lewin, K. (1946). Action research and minority problems. *Journal of Social Issues*, 2: 34-46.

Li, X., Xu, Z. R., Tang, N., Ye, C., Zhu, X. L., Zhou, T., & Zhao, Z. H. (2015). Effect of intervention using a messaging app on compliance and duration of treatment in orthodontic patients. *Clinical Oral Investigations*, *20*(8), 1849–1859. https://doi.org/10.1007/s00784-015-1662-6

Liang, Y., Fan, H. W., Fang, Z., Miao, L., Li, W., Zhang, X., . . . Chen, X. A. (2020). OralCam: Enabling Self-Examination and Awareness of Oral Health Using a Smartphone Camera. *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. https://doi.org/10.1145/3313831.3376238

Lin, C. Y., Peng, K. L., Chen, J., Tsai, J. Y., Tseng, Y. C., Yang, J. R., & Chen, M. H. (2014). Improvements in dental care using a new mobile app with cloud services. *Journal of the Formosan Medical Association*, 113(10), 742–749. https://doi.org/10.1016/j.jfma.2014.02.009

Lingard, L., Albert, M., & Levinson, W. (2008). Grounded theory, mixed methods, and action research. *BMJ, 337*, a567–a567. https://doi.org/10.1136/bmj.39602.690162.47

Listl, S., Galloway, J., Mossey, P. A., & Marcenes, W. (2015). Global Economic Impact of Dental Diseases. *Journal of dental research, 94*(10), 1355–1361. https://doi.org/10.1177/0022034515602879

Liu, X., Faes, L., Kale, A. U., Wagner, S. K., Fu, D. J., Bruynseels, A., . . . Denniston, A. K. (2019). A comparison of deep learning performance against health-care professionals in detecting diseases from medical imaging: a systematic review and meta-analysis. *The Lancet Digital Health, 1*(6), e271–e297. https://doi.org/10.1016/s2589-7500(19)30123-2

Liu, Y., Osvalder, A., & Karlsson, M. (2010). Considering the Importance of User Profiles in Interface Design. *In User Interfaces*. Mátrai, R. (Ed), 61–80.

Ludvigsen, M. (2006). *Designing for social interaction: physical, co-located social computing*. (Doctoral Dissertation) Arkitektskolen Aarhus.

Luger, W. A., & Stubblefield, G. F. (1997). *Artificial Intelligence: Structures and Strategies for Complex Problem Solving* (3rd ed.). Benjamin-Cummings Pub Co.

Luna, D., Peracchio, L. A. and de Juan, M. D. (2008). Cross–cultural and cognitive aspects of website navigation. *Academy of Marketing Journal*, Vol. 30 No. 4, pp. 397-410.

Lykke-Olesen, A. (2006). Space as interface: bridging the gap with cameras. (Doctoral Dissertation) Arkitektskolen Aarhus.

Machado, J. P., Lam, X. T., & Chen, J. W. (2018). Use of a clinical decision support tool for the management of traumatic dental injuries in the primary dentition by novice and expert clinicians. *Dental Traumatology, 34*(2), 120–128. https://doi.org/10.1111/edt.12390

Maisonneuve, P., Amar, S., & Lowenfels, A. (2017). Periodontal disease, edentulism, and pancreatic cancer: a meta-analysis. *Annals of Oncology, 28*(5), 985–995. https://doi.org/10.1093/annonc/mdx019

Malvehy, J., Ginsberg, R., Sampietro-Colom, L., Ficapal, J., Combalia, M., & Svedenhag, P. (2021). New regulation of medical devices in the EU: impact in dermatology. *Journal of the European Academy of Dermatology and Venereology, 36*(3), 360–364. https://doi.org/10.1111/jdv.17830

Margolis, E., & Pauwels, L. (2011). *The SAGE handbook of visual research methods*. Los Angeles: Sage Publications.

Martin, J., & Barnett, J. (2012). Integrating the results of user research into medical device development: insights from a case study. *BMC Medical Informatics and Decision Making, 12*(1), 74. https://doi.org/10.1186/1472-6947-12-74

Martin, R. (1991). School children’s attitudes towards computers as a function of gender, course subjects, and availability of home computers. *Journal of Computer Assisted Learning 7*:187–194.

Martinengo, L., Van Galen, L., Lum, E., Kowalski, M., Subramaniam, M., & Car, J. (2019). Suicide prevention and depression apps' suicide risk assessment and management: a systematic assessment of adherence to clinical guidelines. *BMC Medicine, 17*(1). https://doi.org/10.1186/s12916-019-1461-z

Matsuyama, Y., Aida, J., Tsuboya, T., Hikichi, H., Kondo, K., Kawachi, I., & Osaka, K. (2017). Are Lowered Socioeconomic Circumstances Causally Related to Tooth Loss? A Natural Experiment Involving the 2011 Great East Japan Earthquake. *American journal of epidemiology, 186*(1), 54–62. https://doi.org/10.1093/aje/kwx059

Mattila, K., Nieminen, M., Valtonen, V., Rasi, V., Kesaniemi, Y., Syrjala, S., Jungul, P., Isoluoma, M., Hietaniemi, K., Jokinen, M. & Huttunen, J. (1989). Association between dental health and acute myocardial infarction. *British Medical Journal 298*, 779–782.

McCarney, R., Warner, J., Iliffe, S., van Haselen, R., Griffin, M., & Fisher, P. (2007). The Hawthorne Effect: a randomised, controlled trial. *BMC Medical Research Methodology, 7*(1). https://doi.org/10.1186/1471-2288-7-30

McCarthy, J. J., Minsky, M., Rochester, N., & Shannon, C. E. (2006). A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence, August 31, 1955. *Ai Magazine, 27*(4), 12. https://doi.org/10.1609/aimag.v27i4.1904

McGrenere, J. and Ho, W. (2000): Affordances: Clarifying and Evolving a Concept. *In: Proceedings of Graphics Interface 2000* May 15-17, 2000, Montreal, Quebec, Canada. pp. 179-186

McIntyre, J. D., Lee, J. Y., Trope, M., & Vann, W. F., Jr (2007). Management of avulsed permanent incisors: a comprehensive update. *Pediatric dentistry, 29*(1), 56–63.

McKnight, L. (2010). Designing for ADHD in search of guidelines. *In IDC 2010 Digital Technologies and Marginalized Youth Workshop* (Vol. 30).

McNiff, J., & Whitehead, J. (2002) *Action research: Principles and practice*. (2nd ed.) Routledge Falmer.

Medical Device Coordination Group. (2019). *Guidance on qualification and classification of software in Regulation (EU) 2017/745 - MDR and Regulation (EU) 2017/746 - IVDR*. Retrieved November 10, 2022, from https://ec.europa.eu/docsroom/documents/37581

Megahed, Y. (2017). On research by design. *Arq-architectural Research Quarterly, 21*(4), 338–343. https://doi.org/10.1017/s1359135518000179

Meyer, J. (2000). Qualitative research in health care: Using qualitative methods in health-related action research. *BMJ, 320*(7228), 178–181. https://doi.org/10.1136/bmj.320.7228.178

Meyer, J., & Batehup, L. (1997). Action research in health-care practice: Nature, present concerns and future possibilities. *NT Research, 2*(3), 175–184. https://doi.org/10.1177/174498719700200304

Michaud, D. S., Fu, Z., Shi, J., & Chung, M. (2017). Periodontal Disease, Tooth Loss, and Cancer Risk. *Epidemiologic Reviews, 39*(1), 49–58. https://doi.org/10.1093/epirev/mxx006

Miller, L. A. (1974). Programming by non-programmers. *International Journal of Man-Machine Studies, 6*(2), 237–260. https://doi.org/10.1016/S0020-7373(74)80004-0

Muehlematter, U. J., Daniore, P., & Vokinger, K. N. (2021). Approval of artificial intelligence and machine learning-based medical devices in the USA and Europe (2015–20): a comparative analysis. *The Lancet Digital Health, 3*(3), e195–e203. https://doi.org/10.1016/s2589-7500(20)30292-2

Nadin, M. (1988). Interface design: A semiotic paradigm. *Semiotica, 69*(3–4), 269–302.

Nagendran, M., Chen, Y., Lovejoy, C. A., Gordon, A. C., Komorowski, M., Harvey, H., . . . Maruthappu, M. (2020). Artificial intelligence versus clinicians: systematic review of design, reporting standards, and claims of deep learning studies. *BMJ*, m689. https://doi.org/10.1136/bmj.m689

Nayak, P. P., Nayak, S. S., Sathiyabalan, D., Aditya, N., & Das, P. (2017). Assessing the Feasibility and Effectiveness of an App in Improving Knowledge on Oral Cancer—an Interventional Study. *Journal of Cancer Education, 33*(6), 1250–1254. https://doi.org/10.1007/s13187-017-1239-y

Neumuller, M. (2001). *Hypertext semiotics in the commercialized internet*. (Unpublished Doctoral Dissertation). Wirtschaftsuniversität Wien, Vienna.

Newell, A., & Simon, H. A. (1961). GPS, a program that simulates human thought. In Billing H. (Ed.) *Lernende Automaten*. Munchen: R. Oldenburg.

Newell, A., and Simon, H. A. (1972). *Human Problem Solving*. Englewood Cliffs, New Jersey: Prentice-Hall

NHS Health Scotland (2011). *Childsmile program*. Retrieved October 30, 2022, from https://www.child-smile.org.uk/professionals/about-childsmile.aspx

Nielsen J. (2018). *Ten usability heuristics*. Retrieved June 21, 2018, from https://www.nngroup.com/articles/ten-usability-heuristics/webcite

Nielsen, J. (1994). *10 Usability Heuristics for User Interface Design*. Retrieved October 30, 2022, from http://www.useit.com/articles/ten-usability-heuristics/

Niemiec, E. (2022). Will the EU Medical Device Regulation help to improve the safety and performance of medical AI devices? *DIGITAL HEALTH, 8*. https://doi.org/10.1177/20552076221089079

Nisbett, R. E., Peng, K., Choi, I., & Norenzayan, A. (2001). Culture and systems of thought: holistic versus analytic cognition. *Psychological review*, *108*(2), 291–310. https://doi.org/10.1037/0033-295x.108.2.291

Noiwan, J., & Norcio, A. F. (2006). Cultural differences on attention and perceived usability: Investigating color combinations of animated graphics. *International Journal of Human-computer Studies, 64*(2), 103–122. https://doi.org/10.1016/j.ijhcs.2005.06.004

Nolen, S. L., Giblin-Scanlon, L. J., Boyd, L. D., & Rainchuso, L. (2018). Development and Testing of a Smartphone Application Prototype for Oral Health Promotion. *Journal of dental hygiene: JDH, 92*(2), 6–14.

Norman, D. A. (1988): *The Psychology of Everyday Things*. New York, Basic Books

Norman, D. A. (1999). Affordances, conventions, and design. *Interactions 6*, 3 pp. 38-42.

Norman, D. A. (1998). *The Invisible Computer*. Cambridge: MA, MIT Press.

Norman, K., & Kirakowski, J. (2018). *The Wiley Handbook of Human-Computer Interaction Set* (1st ed.). Wiley-Blackwell.

Offenbacher, S., Katz, V., Fertik, G., Collins, J., Boyd, D., Maynor, G., . . . Beck, J. (1996). Periodontal Infection as a Possible Risk Factor for Preterm Low Birth Weight. *Journal of Periodontology, 67*(10s), 1103–1113. https://doi.org/10.1902/jop.1996.67.10s.1103

OJ L 117, 5.5.2017. *Regulation (EU) 2017/745 of the European Parliament and of the Council of 5 April 2017 on medical devices, amending Directive 2001/83/EC, Regulation (EC) No 178/2002 and Regulation (EC) No 1223/2009 and repealing Council Directives 90/385/EEC and 93/42/EEC*. 2017 May 05. URL: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32017R0745

OJ L 169, 12.7.1993. *Council Directive 93/42/EEC of 14 June 1993 concerning medical devices* URL: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:31993L0042

Osvalder, A-L., and Ulfvengren, P. (2009). Human-technology systems. In Bhgard, G et al. (Eds.) *Work and technology on human terms*, Prevent, Sweden.

Owsley, C., Ball, K., Sloane, M. E., Roenker, D. L., and Bruni, J. R., (1991). Visual/cognitive correlates of vehicle accidents in older drivers. *Psychology and Aging, 6*(3), 403-415.

Palmer, S. (2017). Swipe Right for Health Care: How the State May Decide the Future of the mHealth App Industry in the Wake of FDA Uncertainty. *Journal of Legal Medicine, 37*(1–2), 249–263. https://doi.org/10.1080/01947648.2017.1303289

Panescu, D. (2009). Medical device development. In conference proceedings: Annual International *Conference of the IEEE Engineering in Medicine and*

Biology Society. IEEE Engineering in Medicine and Biology Society. Annual Conference. pp. 5591–5594.

Paxton, A., & Turner, E. T. (1984). The application of human factors to the needs of the novice computer user. *International Journal of Man-machine Studies, 20*(2), 137–156. https://doi.org/10.1016/s0020-7373(84)80014-0

Peirce, C. S. (1931). *Collected Writings*. In Hartshorne, C., Weiss, P. and Burks, A. (Eds.). Vol. 8, Harvard University Press, Cambridge, MA.

Peres, M. A., Macpherson, L. M. D., Weyant, R. J., Daly, B., Venturelli, R., Mathur, M. R., . . . Watt, R. G. (2019). Oral diseases: a global public health challenge. *The Lancet, 394*(10194), 249–260. https://doi.org/10.1016/s0140-6736(19)31146-8

Philbin, M., Meier, E., Huffman, S., and Boverie, P. (1995). A survey of gender and learning styles. *Sex Roles, 32*(7/8), 485-494.

Kraft, P., & Bansler J. P. (1994). The Collective Resource Approach: the Scandinavian experience. *Scandinavian Journal of Information Systems, 6*(1), 71–84. Retrieved from http://iris.cs.aau.dk/tl_files/volumes/volume06/no1/05_bansler_p71-84.pdf

Pithon, M. M., Nascimento, C. C., Barbosa, G. C. G., & Coqueiro, R. D. S. (2014). Do dental esthetics have any influence on finding a job? *American Journal of Orthodontics and Dentofacial Orthopedics, 146*(4), 423–429. https://doi.org/10.1016/j.ajodo.2014.07.001

Preece, J. (2002). *Interaction Design – beyond human-computer interaction*. John Wiley & Sons, Inc., USA.

Privitera, M. (2017). Designing Industrial Design in the Highly Regulated Medical Device Development Process. Defining our valuable contribution towards usability. *The Design Journal, 20*(sup1), S2190–S2206. https://doi.org/10.1080/14606925.2017.1352735

Privitera, M., Southee, D., & Evans, M. (2015). Collaborative design processes in medical device development. In conference proceedings: *11th European Academy of Design Conference* https://doi.org/10.13140/RG.2.1.2711.4967

Pujar, P., & Subbareddy, V. V. (2013). Evaluation of the tooth brushing skills in children aged 6–12 years. *European Archives of Paediatric Dentistry, 14*(4), 213–219. https://doi.org/10.1007/s40368-013-0046-3

Pulijala, Y., Ma, M., Ju, X., Benington, P., & Ayoub, A. (2016). Efficacy of three-dimensional visualization in mobile apps for patient education regarding orthognathic surgery. *International Journal of Oral and Maxillofacial Surgery, 45*(9), 1081–1085. https://doi.org/10.1016/j.ijom.2016.04.002

Pussinen, P. J., Alfthan, G., Rissanen, H., Reunanen, A., Asikainen, S., & Knekt, P. (2004). Antibodies to Periodontal Pathogens and Stroke Risk. *Stroke, 35*(9), 2020–2023. https://doi.org/10.1161/01.str.0000136148.29490.fe

Qari, A. H. (2019). *Assessing the quality of mobile apps for oral health: Content analysis and usability*. (Doctoral Dissertation). Boston University Henry M. Goldman School of Dental Medicine. Retrieved from https://open.bu.edu/handle/2144/37387?show=full

Radnai, M., Gorzó, I., Nagy, E., Urbán, E., Novák, T., & Pál, A. (2004). A possible association between preterm birth and early periodontitis. A pilot study. *Journal of clinical periodontology, 31*(9), 736–741. https://doi.org/10.1111/j.1600-051X.2004.00564.x

Rahaei, Z., Moradian, E., & Falahati-Marvast, F. (2021). Improving dental-oral health learning in students using a mobile application ("My tooth"): A controlled before and after study. *International Journal of Dental Hygiene, 20*(3), 512–518. https://doi.org/10.1111/idh.12504

Reason, P., & Bradbury, H. (2001). Introduction: Inquiry and participation in search of a world worthy of human aspiration. In Reason P, Bradbury H (Eds.) *Handbook of action research: participative inquiry and practice*. Sage Publications.

Republic of Türkiye Ministry of Health General Directorate of Health Information Systems. (2018). *Health Statistics Yearbook 2018*. Retrieved November 20, 2020, from https://sbsgm.saglik.gov.tr/TR-62398/saglik-istatistikleri-yilligi-2018-yayinlandi.html

Ribeiro, M. A. O., Correa, C. G., & Nunes, F. L. S. (2017). Gamification as a Learning Strategy in a Simulation of Dental Anesthesia. 2017 *19th Symposium on Virtual and Augmented Reality (SVR)*. https://doi.org/10.1109/svr.2017.42

Righolt, A., Jevdjevic, M., Marcenes, W., & Listl, S. (2018). Global-, Regional-, and Country-Level Economic Impacts of Dental Diseases in 2015. *Journal of Dental Research, 97*(5), 501–507. https://doi.org/10.1177/0022034517750572

Rodrigez-Andres, D., Juan, M. C., Molla, R., & Mendez-Lopez, M. (2017). A 3D Serious Game for Dental Learning in Higher Education. *2017 IEEE 17th International Conference on Advanced Learning Technologies (ICALT)*. https://doi.org/10.1109/icalt.2017.29

Roethlisberger, F. J., & Dickson, W. J. (1939). *Management and the worker*. Cambridge, MA: Harvard University Press.

Roth, V. J. (2014). The mHealth Conundrum: Smartphones & Mobile Medical Apps – How Much FDA Medical Device Regulation is Required? *North Carolina Journal of Law & Technology, 15*(3), 359–424. Retrieved from https://scholarship.law.unc.edu/cgi/viewcontent.cgi?article=1245&context=ncjolt

Safkan-Seppälä, B., & Ainamo, J. (1992). Periodontal conditions in insulin-dependent diabetes mellitus. *Journal of clinical periodontology, 19*(1), 24–29. https://doi.org/10.1111/j.1600-051x.1992.tb01144.x

Saussure, F. D. (1983), Course in General Linguistics, Duckworth, London (translated by Roy Harris).

Sayar, G., & Kilinc, D. D. (2017). Manual tracing versus smartphone application (app) tracing: a comparative study. *Acta Odontologica Scandinavica, 75*(8), 588–594. https://doi.org/10.1080/00016357.2017.1364420

Scannapieco, F. A., Bush, R. B., & Paju, S. (2003). Associations between periodontal disease and risk for nosocomial bacterial pneumonia and chronic obstructive pulmonary disease. A systematic review. *Annals of periodontology, 8*(1), 54–69. https://doi.org/10.1902/annals.2003.8.1.54

Scapin, D. L. (1981). Computer Commands in Restricted Natural Language: Some Aspects of Memory and Experience. *Human Factors, 23*(3), 365–375. https://doi.org/10.1177/001872088102300311

Schaeffer, N. E. (2012). The Role of Human Factors in the Design and Development of an Insulin Pump. *Journal of Diabetes Science and Technology, 6*(2), 260–264. https://doi.org/10.1177/193229681200600208

Scheerman, J. F. M., van Meijel, B., van Empelen, P., Kramer, G. J. C., Verrips, G. H. W., Pakpour, A. H., . . . van Loveren, C. (2018). Study protocol of a randomized controlled trial to test the effect of a smartphone application on oral-health behavior and oral hygiene in adolescents with fixed orthodontic appliances. *BMC Oral Health, 18*(1). https://doi.org/10.1186/s12903-018-0475-9

Scheibe, M., Reichelt, J., Bellmann, M., & Kirch, W. (2015). Acceptance Factors of Mobile Apps for Diabetes by Patients Aged 50 or Older: A Qualitative Study. *Medicine 2.0, 4*(1), e1. https://doi.org/10.2196/med20.3912

Schon, D. A. (1984). *The Reflective Practitioner* (1st Edition). Basic Books.

Scott, K. M., Gome, G. A., Richards, D., & Caldwell, P. (2015). How trustworthy are apps for maternal and child health? *Health and Technology*, *4*(4), 329–336. https://doi.org/10.1007/s12553-015-0099-x

Sedgwick, P., & Greenwood, N. (2015). Understanding the Hawthorne effect. *BMJ*, h4672. https://doi.org/10.1136/bmj.h4672

Sharif, M. O., & Alkadhimi, A. (2019). Patient-focused oral hygiene apps: an assessment of quality (using MARS) and knowledge content. *British Dental Journal, 227*(5), 383–386. https://doi.org/10.1038/s41415-019-0665-0

Shen, J., Zhang, C. J. P., Jiang, B., Chen, J., Song, J., Liu, Z., . . . Ming, W. (2019). Artificial intelligence versus Clinicians in Disease Diagnosis: Systematic review. *JMIR Medical Informatics*, *7*(3), e10010. https://doi.org/10.2196/10010

Sherman, R. C., End, C., Kraan, E., Cole, A., Campbell, J., Birchmeier, Z., and Klausner, J. (2000). The Internet gender gap among college students: forgotten but not gone? *CyberPsychology & Behavior 3*: 885–894. https://doi.org/10.1089/10949310050191854

Shlossman, M., Knowler, W. C., Pettitt, D. J., & Genco, R. J. (1990). Type 2 diabetes mellitus and periodontal disease. *Journal of the American Dental Association* (1939), 121(4), 532–536. https://doi.org/10.14219/jada.archive.1990.0211

Shneiderman, B. (1976). Exploratory experiments in programmer behavior. *International Journal of Computer & Information Sciences, 5*(2), 123–143. https://doi.org/10.1007/bf00975629

Simon, D. P. & Simon, H. A. (1978). Individual differences in solving physics problems. In Siegler, R. (Ed.). *Children's Thinking: What Develops*. Hillsdale, New Jersey: Erlbaum.

Simons, H. (1981) Process evaluation in schools. In: Lacey, C., Lawton, D. (Ed.) *Issues in Evaluation and Accountability*. London: Methuen; pp115-144.

Smith, A., & Yetim, F. (2004). Global human–computer systems: cultural determinants of usability. *Interacting With Computers, 16*(1), 1–5. https://doi.org/10.1016/j.intcom.2003.11.001

Spinuzzi, C. (2005). The Methodology of Participatory Design. *Technical Communication, 52*(2), 163–174. Retrieved from https://repositories.lib.utexas.edu/bitstream/2152/28277/2/SpinuzziTheMethodologyOfParticipatoryDesign.pdf

Statista (2022). *Google Play: number of available medical apps as of Q2 2022*. Retrieved November 7, 2022, from https://www.statista.com/statistics/779919/health-apps-available-google-play-worldwide/

Stein, C. D., Xiao, X., Levine, S., Schleyer, T. K., Hochheiser, H., & Thyvalikakath, T. P. (2016). A prototype mobile application for triaging dental emergencies. *The Journal of the American Dental Association, 147*(10), 782-791.e1. https://doi.org/10.1016/j.adaj.2016.03.021

Stephanidis, C., Paramythis, A., Sfyrakis, M., Stergiou, A., Maou, N., Leventis, A.… Karagiannidis, C. (1998) Adaptable and adaptive user interfaces for disabled users in the AVANTI Project. In S. Triglia, A. Mullery, M. Campolargo, H. Vanderstraeten, & M. Mampaey (Eds.). *Intelligence in services and networks: Technology for ubiquitous telecom services*. Berlin: Springer.

Sternberg, R. J. (1998). Abilities are forms of developing expertise. *Educational Researcher, 27*, 11–20.

Sternberg, R. J. (2008). Schools should nurture wisdom. In B. Z. Presseisen (Ed.), *Teaching for intelligence* (2nd ed., pp. 61–88). Thousand Oaks, CA: Corwin.

Stoyanov, S. R., Hides, L., Kavanagh, D. J., Zelenko, O., Tjondronegoro, D., & Mani, M. (2015). Mobile App Rating Scale: A New Tool for Assessing the Quality of Health Mobile Apps. *JMIR mHealth and uHealth, 3*(1), e27. https://doi.org/10.2196/mhealth.3422

Susman, G. I., & Evered, R. D. (1978). An Assessment of the Scientific Merits of Action Research. *Administrative Science Quarterly, 23*(4), 582. https://doi.org/10.2307/2392581

Taylor, P., Manzini, E., & Rizzo, F. (2011). Small projects / large changes: Participatory design as an open participated process. *CoDesign: International Journal of CoCreation in Design and the Arts* (May 2012), 37–41.

Tekinarslan, E. (2008). Computer anxiety: A cross-cultural comparative study of Dutch and Turkish university students. *Computers in Human Behavior, 24*(4), 1572–1584. https://doi.org/10.1016/j.chb.2007.05.011

Thimbleby, H. (2018). Interactive critical systems and how to build them. (K. L. Norman & J. Kirakowski, Eds.), *The Wiley Handbook of Human-Computer Interaction* (1st ed., Vol. 1, pp. 9–32). John Wiley & Sons Ltd.

Thorstensson, H., & Hugoson, A. (1993). Periodontal disease experience in adult long-duration insulin-dependent diabetics. *Journal of clinical periodontology, 20*(5), 352–358. https://doi.org/10.1111/j.1600-051x.1993.tb00372.x

Throckmorton, A. L., Patel-Raman, S. M., Fox, C. S., & Bass, E. J. (2016). Beyond the VAD: Human Factors Engineering for Mechanically Assisted Circulation in the 21st Century. *Artificial Organs, 40*(6), 539–548. https://doi.org/10.1111/aor.12600

Tiffany, B., Blasi, P., Catz, S. L., & McClure, J. B. (2018). Mobile Apps for Oral Health Promotion: Content Review and Heuristic Usability Analysis. *JMIR mHealth and uHealth, 6*(9), e11432. https://doi.org/10.2196/11432

Toniazzo, M. P., Nodari, D., Muniz, F. W. M. G., & Weidlich, P. (2019). Effect of mHealth in improving oral hygiene: A systematic review with meta-analysis. *Journal of Clinical Periodontology, 46*(3), 297–309. https://doi.org/10.1111/jcpe.13083

Torenvliet, G. (2003). We can't afford it!: the devaluation of a usability term. *In Interactions, 10* (4) pp. 12-17

Torous, J., & Firth, J. (2016). The digital placebo effect: mobile mental health meets clinical psychiatry. *The Lancet Psychiatry, 3*(2), 100–102. https://doi.org/10.1016/s2215-0366(15)00565-9

Torous, J., Stern, A. D., & Bourgeois, F. T. (2022). Regulatory considerations to keep pace with innovation in digital health products. *Npj Digital Medicine, 5*(1). https://doi.org/10.1038/s41746-022-00668-9

Turing, A. M. (1950). Computing Machinery and Intelligence. *Mind, 59*(236), pp. 433–460

Turner, A. (2022). *How Many People Have Smartphones Worldwide*. Retrieved October 5, 2022, from https://www.bankmycell.com/blog/how-many-phones-are-in-the-world

U.S. Food and Drug Administration, Center for Devices and Radiological Health, & Center for Biologics Evaluation and Research. (2022). *Policy for device software functions and mobile medical applications*. Retrieved November 8, 2022, from https://www.fda.gov/regulatory-information/search-fda-guidance-documents/policy-device-software-functions-and-mobile-medical-applications

U.S. Food and Drug Administration (2020). *Unique Device Identification System (UDI System)*. Retrieved November 10, 2022, from https://www.fda.gov/medical-devices/device-advice-comprehensive-regulatory-assistance/unique-device-identification-system-udi-system

U.S. Food and Drug Administration (2021). *Artificial intelligence and machine learning in software as a medical device*. Retrieved November 13, 2022, from https://www.fda.gov/medical-devices/software-medical-device-samd/artificial-intelligence-and-machine-learning-software-medical-device

Ulrich, K. & Eppinger, S. (2008). *Product Design and Development*, New York: McGraw-Hill Higher Education.

Umino, M. & Nagao, M. (1993). Systemic diseases in elderly dental patients. *International Dental Journal 43*, 213–218.

Underwood, B., Birdsall, J., & Kay, E. (2015). The use of a mobile app to motivate evidence-based oral hygiene behaviour. *British Dental Journal, 219*(4), E2–E2. https://doi.org/10.1038/sj.bdj.2015.660

Uthoff, R. D., Song, B., Sunny, S., Patrick, S., Suresh, A., Kolur, T., . . . Liang, R. (2018). Point-of-care, smartphone-based, dual-modality, dual-view, oral cancer screening device with neural network classification for low-resource communities. *PLOS ONE, 13*(12), e0207493. https://doi.org/10.1371/journal.pone.0207493

van Deursen, A., & van Diepen, S. (2013). Information and strategic Internet skills of secondary students: A performance test. *Computers & Education, 63*, 218–226. https://doi.org/10.1016/j.compedu.2012.12.007

van Deursen, A., & van Dijk, J. (2009). Improving digital skills for the use of online public information and services. *Government Information Quarterly, 26*(2), 333–340. https://doi.org/10.1016/j.giq.2008.11.002

van Norman, G. A. (2016). Drugs and Devices. *JACC: Basic to Translational Science, 1*(5), 399–412. https://doi.org/10.1016/j.jacbts.2016.06.003

Vatrapu, R., & Pérez-Quiñones, M. A. (2006). Culture and usability evaluation: the effects of culture in structured interviews. *Journal of Usability Studies Archive, 1*(4), 156–170.

Verganti, R. (2009). Design-driven Innovation: Changing the Rules of Competition by Radically Innovating What Things Mean. Boston, Massachusetts: Harvard Business Press.

Wasylewicz, A. T. M., & Scheepers-Hoeks, A. M. J. W. (2019). Clinical decision support systems. In P. L. Kubben, M. Dumontier, & A. Dekker (Eds.), *Fundamentals of Clinical Data Science* (1st ed., pp. 153–170). Springer. https://doi.org/10.1007/978-3-319-99713-1

Webb C. (1989). Action research: philosophy, methods and personal experiences. *Journal of advanced nursing*, *14*(5), 403–410. https://doi.org/10.1111/j.1365-2648.1989.tb01548.x

Welford, A. T. (1977). Motor performance, In: Birren, J.E., Schaie, K.W. (Eds.), *Handbook of the psychology of aging*. Van Nostrand Reinhold Co., New York.

Whitehead, L., & Seaton, P. (2016). The Effectiveness of Self-Management Mobile Phone and Tablet Apps in Long-term Condition Management: A Systematic Review. *Journal of medical Internet research*, *18*(5), e97. https://doi.org/10.2196/jmir.4883

World Health Organization (WHO) (2016). *Social Determinants of Health*. Retrieved December 19, 2022, from http://www.who.int/social_determinants/sdh_definition/en/

World Health Organization (WHO) and International Telecommunication Union (ITU) (2021). *Mobile technologies for oral health: an implementation guide*. Geneva. License: CC BY-NC-SA 3.0 IGO.

World Health Organization (WHO) (2016). *Diagnostic errors*. World Health Organization. License: CC BY-NC-SA 3.0 IGO. Retrieved from https://apps.who.int/iris/handle/10665/252410

World Health Organization (2022). *Oral health*. Retrieved October 24, 2022, from https://www.who.int/news-room/fact-sheets/detail/oral-health

Xiao, J., Fiscella, K. A., & Meyerowitz, C. (2021). mDentistry. *The Journal of the American Dental Association, 152*(9), 713–716. https://doi.org/10.1016/j.adaj.2021.06.004

Xiao, J., Luo, J., Ly-Mapes, O., Wu, T. T., Dye, T. D., Jallad, N. A., . . . Fiscella, K. (2021). Assessing a smartphone app (AICaries) that uses artificial intelligence to detect dental caries in children and provides interactive oral health education: Protocol for a design and usability testing study. *JMIR Research Protocols*, *10*(10), e32921. https://doi.org/10.2196/32921

Yin, A. L., Hachuel, D., Pollak, J. P., Scherl, E. J., & Estrin, D. (2019). Digital Health Apps in the Clinical Care of Inflammatory Bowel Disease: Scoping Review. *Journal of Medical Internet Research, 21*(8), e14630. https://doi.org/10.2196/14630

Zaror, C., Espinoza-Espinoza, G., Atala-Acevedo, C., Muñoz-Millán, P., Li, Y., Clarke, K., . . . Mariño, R. (2018). Validation and usability of a mobile phone application for epidemiological surveillance of traumatic dental injuries. *Dental Traumatology, 35*(1), 33–40. https://doi.org/10.1111/edt.12444

Zicari, R. V., Ahmed, S., Amann, J., Braun, S. A., Brodersen, J., Bruneault, F., . . . Wurth, R. (2021). Co-Design of a Trustworthy AI System in Healthcare: Deep Learning Based Skin Lesion Classifier. *Frontiers in Human Dynamics, 3*. https://doi.org/10.3389/fhumd.2021.688152

Zimmerman, J., Stolterman, and E., Forlizzi, J. (2010). An analysis and critique of research through design: towards a formalisation of research approach. *DIS '10: Proceedings of the 8th ACM Conference on Designing Interactive Systems*. August 16-20, Aarhus, Denmark. https://doi.org/10.1145/1858171.1858228

Zolfaghari, M., Shirmohammadi, M., Shahhosseini, H., Mokhtaran, M., & Mohebbi, S. Z. (2021). Development and evaluation of a gamified smartphone mobile health application for oral health promotion in early childhood: a randomized controlled trial. *BMC Oral Health, 21*(1). https://doi.org/10.1186/s12903-020-01374-2

Zotti, F., Dalessandri, D., Salgarello, S., Piancino, M., Bonetti, S., Visconti, L., & Paganelli, C. (2015). Usefulness of an app in improving oral hygiene compliance in adolescent orthodontic patients. *The Angle Orthodontist, 86*(1), 101–107. https://doi.org/10.2319/010915-19.1

Url-1 <https://commons.wikimedia.org/wiki/File:%EB%B2%95%EB%9E%91%EC%A7%88%ED%98%95%EC%84%B1%EB%B6%80%EC%A0%84%EC%A6%9D1.jpg>, date retrieved 05.01.2022.

Url-2 <https://dentistchannel.online/enamel-hypoplasia-a-disease-effecting-enamel-of-teeth/article?for=people>, date retrieved 05.01.2022.

Url-3 <https://webuildsmileappeal.com/instructions/tongue-thrust-correction/>, date retrieved 05.01.2022.

Url-4 <https://www.smileworksliverpool.co.uk/blog/deep-overbite-treatment-adults/>, date retrieved 05.01.2022.

Url-5 <https://www.practo.com/healthfeed/does-teeth-cleaning-damage-teeth-39823/post>, date retrieved 05.01.2022.

Url-6 <https://encyclopedia.pub/entry/30494>, date retrieved 05.01.2022.

Url-7 <https://blog.sesamehub.com/nikolovski-john96/files/2013/01/underbite-2.jpg>, date retrieved 05.01.2022.

Url-8 <https://www.muhadharaty.com/files/lectures/017/file17534.pptx_d/image2.png.jpg>, date retrieved 05.01.2022.

APPENDICES

Appendix – A: Original Versions of the Participants' Comments

A) Participant comments in the research 1

1) The comments on (a) *Amelogenesis Imperfecta* and (b) *Trauma-related Enamel Hypoplasia:*

P1: (a) Dişlerin yüzeyleri tırtık tırtık, kanal kanal olmuş, kol kol akarsu gibi, renginde de sarılaşma kahverengileşme var tüm diş yüzeyinde" (b) "Diş üst yarısı normal alt yarısı anormal, girdili çıktılı, krater krater olmuş ve sararma var."

P2: (b) "Diş, plastiği ısıtırlar ya, plastik ısınınca erimiş gibi gözüküyor. Diş, olması gereken geometrinin dışına çıkmış."

P3: (a) "Dişlerin tamamının yüzeyinde deformasyon var… Doğuştan mine zedelenmesi derdim."

P4: (a) "…Dişlerin üzerinde krater görüntüsü var… Bütünlük şekil bozukluğu hatta yüzeyde pürüzlülük en uygunu."

2) The comments on *Open Bite*:

P1: "Dişlerin arasında ayrıklık var. Madonna gibi, sadece ayrıklığın yeri farklı. (Açık kapanış senin için bir şey ifade ediyor mu?) Hayır."

P2: "Çocuğun çenesi tam basmıyor. Ön dişler havada kalıyor. (Açık kapanış bir şey ifade ediyor mu?) Hayır etmiyor."

P3: "Çok simetrik bir ayrılma var dişler arasında. Diş arası boşluk. Simetrik bir şekilde diş arası boşluk üst ve alt dişlerin arasında. (Açık kapanış bir şey ifade eder miydi?) Hayır etmezdi."

P4: "Dişler tam kapanmıyor. Hatta aralarında boşluk var. (Açık kapanış bir şey ifade eder mi?) Hayır çok anlaşılmıyor."

P5: "Dişler tam kapanmıyor. Açık kapanış benim için de bir şey ifade etmezdi. Sanki yanlardan basık, önden çıkık."

3) The comments on *Deep Bite*:

P1: “Dişlerin yönü yanlış. Dişler arkaya yatık çıkmış, eğer eksiklik ve boyutta bir farklılık yoksa böyle. (Derin kapanış ifade eder mi?) bir tık evet bu vakada. Burada diş üstünü örtmüş gibi ondan anlamlı.”

P2: “Alt çene orta dişleri çıkmamış. (Dişlerin var olduğunu belirttim) Çene kapalıyken alt çene orta dişler gözükmüyor. (Üst dişler alt dişleri örtmüş denilebilir mi?) evet denilebilir. (Derin kapanış bir şey ifade ediyor mu?) Hayır. Sanırım daha açıklayıcı bir şeyler lazım. Bir iki kelimelik değil de daha uzun tanımlara ihtiyaç var.”

P3: “Üst dişler alt dişlerin üzerine binmesi, alt dişlerin görünmemesi. (Derin kapanış bir şey ifade eder miydi?) Şu an bunu göstererek sorunca bir kısım ifade etti.”

P4: “Alt dişler yok sanki üstüne mi kapanmış gibi? (Dişler var diyorum) Alt tarafta kapandığında alt dişler gözükmüyor. Ha bakardım alt dişler gözükmüyor işaretlerdim ama derin kapanış bir şey ifade etmiyor.”

P5: “Kapanış bozukluğu, dişler üst üste gelmiş. (Derin kapanış bir şey ifade eder mi?) Hayır söylemezdim. Kapanış bozukluğu var derdim.”

4) The comments on *Gingival Recession*:

P1: “Dişetinde çekilme var. Diş her an düşecek gibi. Köklerine doğru dişeti uzamış gibi. Et açılmış dişin dibi gözüküyor. Dişeti açılması olabilir. (Dişeti çekilmesini işaretler misin diye sorduğumda) Dişeti çekilmesi deyince sanki bütün dişlerde az da olsa görülen çekilme gibi, burada daha garip bir durum var. Çok aşırı çekilmiş. Sanki burada 1 mm gibi bir ölçü olsa daha rahat söylerdim.”

P2: “Dişinin kökü çıkmış, etten kökü gözüküyor. (Dişeti çekilmesi senin için uygun bir tabir olur mu?) Dişeti çekilmesi birden çok et çekilmesi olduğunda derdim. Et çekilirse hepsinde çekilirdi diye düşünürdüm. Sanki kök açığa çıkmış gibi.”

P3: “Dişin kökünün görünmesi. Dişeti kaybı. (Dişeti çekilmesi uygun olur muydu?) Evet işaretlerdim.”

5) The comments on *Geograpic Tongue*:

P2: “Diline uzaylılar inmiş. İşaretler filmindeki gibi. Dilde hem kırmızı noktalar hem de beyaz helezonik desenler oluşmuş derdim. (Papillerden bahsettim ve alanlar arasında bir fark var mı diye sordum) Kendi dilime bakardım. Hangisi olması gereken bilmiyorum ama ikisi birden olmamalı. (Dilde düzleşme mantıklı gelir miydi?) Hayır kullanmazdım. Dilde bazı desenler oluşmuş derdim. Hangisinin doğru olduğunu bilmediğim için doğru yorum yapamazdım sanırım. Ama bu kadar heterojen olması da problem derdim.”

P3: “Dilde trippy desenler. Dilde beyaz desenler. Dilde noktacık kaybı. Dil ucunda pürüzsüzleşme ve düzleşme. O pürüzler gitmiş.”

P4: “Kırmızı noktalar değil de bu beyaz kısımlar önemli sanki burada. Normalde dilin yüzeyinde bir doku olması lazım sanki o kaybolmuş gibi. Dilde düzleşmeden ne demek istediğini anladım ama bence desen daha mantıklı. Dilin yüzeyi aslında tüycük gibi yapılar var ve onlar kaybolmuş.”

6) The comments on *Anterior and Posterior Crossbite*:

P1: “Üst dişler bu sefer de içe doğru girik, alt dişlerin arkasına doğru düşmüş halbuki tam tersi olması lazım… Dişlerin yönü yada çeneyle ilgili bir sıkıntı var gibi. Üst dişler alt dişlerin arkasında çıkmış gibi. Üst ve alt dişler arasında boşluk var. Çeneyle mi ilgili, konumuyla ilgili yoksa dişlerin boyutuyla ilgili mi bilemiyorum.”

P2: “Üst çene alt çenenin iç kısmında kalıyor bu da içe kapanış mı?... Köpek dişlerden sonrası askıda kalıyor. Çene birbirine basmıyor. (Arkadaki dişler hakkında ne düşünüyorsun?) Şimdi sen öyle deyince onlar biraz içerde gibi geldi. Doğru mu?”

P3: “Alt çene önde. Alt çenenin üst dişleri kapatması. (Ters kapanış bir şey ifade eder miydi?) Evet ifade ederdi. Çene eklenirse daha açıklayıcı olabilir. Çenede ters kapanma dersen anlardım.” “Kısa üst dişler, üst dört dişte kısalık. (Arka bölgelerde bir şey var mı?) Rengi dokusu normal olduğu için farklı gelmedi. Sayısı mı farklı? (Hayır. Çene ilişkisinde bir problem var mı?) Hayır bilemedim.”

P4: “Kapanışta üst çene içerde kalıyor. Kapandığında üst dişler gözüküyor mu? Tam gözükmüyor.” “Sağlıklı gibi bu. (Arka dişlerle alakalı ne düşünüyorsun?) Sağlıklısında biraz daha dışarı çıkması gerekirken sanki bunlar üst üste binmişler.”

P5: “Deminkinin tam tersi sanki derin kapanış değil tam tersi, alt dişler üst dişleri kapatıyor derdim ben buna, üst çene geride gibi.” “Arka dişler tam oturmuyor. Üst dişler sanki içeri doğru.”

7) The comment of P2 on *Abrasion*:

P2: “Normalde diş etinin altındaki kısmı bilmediğim için, orada madde kaybı var demek aklımın ucundan bile geçmezdi.”

8) The comment of P4 on the methodology of the study:

P4: “Burada sadece resme bakıyoruz aynı deneyim değil. Çocuk geliyor ağrım var ya da kanamam var diyor. Bakıp görüyoruz. Resme bakmak ile aynı şey değil.”

B) Participant comments in the Research 2

1) The comments on the prototypes:

P4: “İlk prototipi tercih ederim. Diğeri çok uzun geldi. Çok uzun. Uzun yazıları okuması ve anlaması zor. Kendi adıma resimlere bakarak daha kolay bir seçim yapabilirim. Burada mesela çapraşıklık var. Sadece resme bakarak var ya da yok diyebilirim. Ama diğer semptomlarda zor çünkü bazılarında çok uzun açıklamalar var.”

P5: “İlk prototip daha çok fikir veriyor ve daha hızlı.”

P4: “Her bir semptom için ayrı bir sayfa var. Halbuki birinci prototipte bütün semptomları alt alta sıralamışsın. Çok kolay bir şekilde tık tık seçiyorsun.”

P2: “Soruları peş peşe alınca daralırdım gibi. 13 aşamadan oluşan bir akış yerine 4 aşama mı olsaydı acaba?” “Mümkün olan en kısa. Seçenekler ile görsellerin aynı yerde olduğu bir akış. Ben mesela bir dakikamı ayırdım bana kaba taslak bir şey çıkardı. Bana şey dese, sen bana bir dakikayı verdin sana bu verileri çıkardım; 5 dakikayı versen daha detaylı bir sonuç çıkarırım dese, ben daha motive olurdum herhalde."

P3: "Bence ihtiyaca göre ikisi de. Çok spesifik ben biliyorsam ne olduğunu bana var mı yok mu diye sormasındansa burada işte söylemiş ağrı salya akıntısı vs. tıklarım. İkincisi daha detaylı hissettirdi. Benim ağız çene sağlığı hakkında bir bilgim olmadığı için, bu daha çok güven verdi. Eğer ihtiyacım olmadığını düşünüyorsam, ay ne uzun sürdü diyebilirim. Ama işte çene sağlığı üzerine bir şey yapıyorum, hekime de gitmeden AI bir şeyler soracak ki ben de cevap vereceğim… Sanırım ikinci, birilerinin bana soru sorması ve yönlendirmesi, benim kendi kendime bulmaya çalışmamdan daha yardımcı. Öteki türlü sanki hepsini kontrol etmem gerekiyormuş gibi geldi. Bana sorular sorulması o anlamda daha iyi geldi. Bir de defa bilmiyorsun, konu ağız sağlığı. Yeterli veriyi verebildim mi kaygısı yaşıyorsun. Ama P2 de soru cevap şeklinde ilerlediği için bu kaygıyı o kadar hissetmezsin."

2) The story of P4:

P4: "Bu arada sadece çocuk değil yetişkinler için de önemli. Bizim şirkete 20 yıl önce bir diş hekimi geldi. Böyle kameralar kurdu. Dediler ki, diş hekimi geldi ücretsiz muayene edilecek isteyen olabilir. Ben de olmak istedim. Bu arada hiçbir şikâyetim yok dişlerimden. Adam kamerayla gösteriyor. Dişimdeki en küçük siyahlığı bile gösteriyor. Bana 12 tane dolgu yapılacak dedi. Ben beynimden vurulmuşa döndüm. Aman allahım dedim. 12 dişte çürük var diyor dolgu yapılacak diyor. Benim zaten kaç tane dişim var? Ondan sonra çok üzüldüm, çok korktum. Halbuki dişimde hiçbir sorunum yok. İyidir benim dişlerim. Sağlıklıdır. Sonra dedim ki bu böyle olmaz. Başka bir hekime gittim. O da amma da abartmış dedi. Her siyahlığa dolgu mu yapılır dedi. 4 tane dolgu çıkarttı. Ben tabi 12'den sonra 4'e ok oldum. Şimdi gittiğimiz diş hekimi dedi ki neden dolgu yapıldı dişinize? Anlattım. Ama dedi bu şekildeki dişlere dolgu yapılmaz ki, sizin dişlerinizde hiçbir şey yok dedi. Bazı siyahlıklar olabilir. Yıllarca onlar kalır. Arada gelirsin kontrol ettirirsin. İlerleme görülürse o zaman dolgu yapılır."

C) Participant comments in the Research 3

The comments on medical content:

- The comments related to medical content in the article 1.a:

Parent 2: "Chornoic villus sampling'i hiç duymadım."

Designer 1: "Enfeksiyöz her ebeveyn bilir mi? Burada bir açıklamaya ihtiyaç olabilir."

Designer 2: "Uyku zamanı ne demek? Uyurken mi yoksa uyku öncesi mi?"

Designer 3: "Bu VKİ ne? Belki yanında i olabilir info için."

Parent 3: "Bu sadece biberon mu? Suluk ya da emzik gibi şeyler dahil değil mi? Emzik var ise evet diyeceğim ama biberonsa hayır diyeceğim."

- The comments related to medical content in the article 1.b:

Parent 1: “Reflü sistemik hastalık sayılmaz herhâlde. Reflü sistemik hastalığa girer mi?”

Parent 2: “Sistemik bir hastalık ne demek? Ben bilmiyorum. Reflü bir sistemik hastalık mı?”

Designer 2: “Sistemikten kasıt ne? Burada sistemik benimle aynı dili konuşmuyor.”

Dentist 1: “Çocuğun alerjisi sistemik bir durum mu?... Sistemik hastalıklar yerine genel sağlığı ile ilişkili bir durum ya da hastalık olarak yazabilirsin.”

Dentist 2: “Reflü sistemik hastalık mıdır? Diz ağrısı sistemik midir? Sistemik değildir, ortopedik bir hastalıktır. Sistemik hastalık anemi olabilir. Çocuğunuzun hastalığı nedir? Daha doğru bir soru olurdu.”

Dentist 3: “Sistemik hastalık kısmı için, bildiğiniz bir hastalığı var mı daha mantıklı olurdu.”

Dentist 1: “Aktif çürüğü anlamayabilir. Tedavi olmamış çürükle dolguyu ayırt etmek istiyorsun ama bunu başka bir şekilde ifade etmek lazım.”

Dentist 3: “Aktif çürüğü anlamayabilir. "Şu anda ağzınızda tedavi edilmemiş çürük diş var mı?” olabilir.”

- The comments related to medical content in the article 1.c:

Designer 2: “İlaç deyince, normalde ilacın ismini arardım.”

Parent 1: “Tetrasiklin için bana bir ticari isim söyler misin? Ben de öğreneyim.”

Parent 2: “Thalidomide ne demek bilmiyorum. Ne mesela bu?”

Parent 3: “Tetrasiklin bu ilacın bileşeni mi? Ancak dikkatli bir ebeveyn bilebilir. Zaten doktor reçete ediyor ben bilemezdim muhtemelen. Bir de bilmiyorsam ne olacak? Burada böyle bir seçenek yok. Ayrıca çocuğu çocuk doktoruna götürüyorsun, o da uygun olan bir ilacı yazıyor diye düşünürüm yani sorgulamam.”

Desginer 1: “Tetrasiklin ne? Türev ne? Herhalde bunların ikisi de bir antibiyotik türü. Hanım hanım! Bizim çocuk tetrasiklin türevi bir antibiyotik kullandı mı? Neyi? Türevi türevi… Burada halk ağzı bir açıklamaya ihtiyaç var.”

Designer 2: “Tetrasiklin türevi antibiyotik çok teknik.”

- The comments related to medical content in the article 1.d:

Dentist 3: “Tetrasiklin hangi hastalıklarında kullanılıyor? Ebeveyn bilemeyebilir. Ticari isimleri yazmak gerekebilir.”

Me: “Ben bu parametreleri literatüde kullanılıyor diye ekledim. Ama belki de artık tetrasiklin kullanılmıyordur. Doktorlar yazmıyor olabilir gebe ve çocuklara.”

Dentist 3: “Ama belki evde vardır, kafasına göre kullanmış olabilir.” “Tetrasiklin sorusunu herkese sormaktan ziyade sadece renklenmeyle ilgili semptomu olan insanlara sormak daha mantıklı olabilir. Çünkü çok nadir olacak muhtemelen bu.”

- The comments related to medical content in the article 2.c:

Dentist 2: “Dudak damak yarığında doğumdan sonra bir cerrahi oluyor ya onu ekleyecek miyim buraya?”

- The comments related to medical content in the article 2.d:

Designer 3: “Burada travmayı soruyor, travmanın ne olduğunu da sorması gerekmez miydi?”

Designer 2: “Çocuğun travmasını soruyor ama hikayeyi istemiyor tarihi istiyor. Bu biraz garip geldi. Bir de hatırlaması zor. Tam tarihi istiyor. O an formlarda bu bilgiler kullanıcıyı kilitliyor. Bu bilgi şu an yok bende bunu öğrendiğimde devam etmek isterim. Doldurmadan geçmek istemem.”

- The comments related to medical content in the article 2.e:

Dentist 3: “Hmm bizim çocuğun reflüsü vardı ama bu listede yok. O zaman ekleye mide koruyucu yazıp geçeceğim çünkü reflüsü var.”

Dentist 2: “Listenin altında neden bir metin kutusu var?”

Me: “Tedaviler çok çeşitli olabilir dolayısıyla bunu algoritmanın anlamayacağı ama kaydedeceği şekilde string olarak alıyorum.”

Dentist 2: “Bazı şeyleri çıkarmak mantıklı olabilir. Her bilgiyi alacağım diye major şeyleri gözden kaçırmak istemezsin. Tedaviler bu kadar önemli mi?”

Me: “Kemoterapi ya da radyoterapi alan çocuklarda oral mucositis olasılığı en üste çıkıyor. Bu atlamak da istemiyorum.”

Dentist 2: “Çocuğu kemoterapi gören ebeveynler bunu bilmezler mi?”

Me: “Ama ben algoritmayı buna göre düzenlemedim işte. Ayrıca doktorlar da ağız hastalıklarından çok anlamıyor. Algoritma doktorlar için de kullanılabilir.”

Dentist 2: “Doğru, ben hala bunu klinikte kullanabileceğim bir şey olarak görüyorum.”

- The comments related to medical content in the article 3.a:

Parent 1: “Çevre kirliliğine maruz kaldı mı? Valla Çin’de yaşasam evet derim ama biz kalıyor muyuzdur? Hayır dedim bilmiyorum olup olmadığını.”

Parent 2: “Çevre kirliliğine maruz kaldı mı çok geniş bir soru. Ben mesela çevre kirliliğine maruz kalıyorum ama Zonguldak’ta yaşamıyorum. Yine de evet diyeceğim.”

Dentist 1: “Çevre kirliliği sorusuna da nasıl cevap vermek lazım? Yani Ankara’da belki hava kirliliği azdır ama çevre bölgelerde fazla olabilir.”

Dentist 2: “Çevre kirliliği var mı? Bu soru çok genel. İzmir’de yaşayan biri çevre kirliliğine maruz kalıyordur.”

Designer 2: “Hava kirliliğini çevre kirliliği olarak algılamayabilir. Etrafta çöp görüyor muyum? Hayır görmüyorum dolayısıyla çevre kirliliği yok diyebilir.”

- The comments related to medical content in the article 3.b:

Parent 2: “Sıklıkla tüketiyor ne demek? Bunu spesifik hale getirmek lazım.”

- The comments related to medical content in the article 3.c:

Designer 1: “Çocuğunuz anne sütüyle kaç ay beslendi? Süreler bana garip geldi. Teknik olarak doğru olabilir ama pratik olarak yanlış. Bu 364 gün olsa 12 ay olmayacak mı? Anlıyorum literatürde böyle geçebileceğini ama bu şekilde alacağın veriye güvenmemeni tavsiye ederim. Güvenilir veri alamıyorsan çıkar bu soruyu.”

- The comments related to medical content in the article 4:

Parent 1: “Çocuğuna sürekli şeker veren bir anne baba bu uygulamayı kullanır mı acaba?”

Dentist 1: “Ailede stres sorusuna Türkiye’de hayır diyecek biri olduğunu sanmıyorum.”

Dentist 2: “Bu her şeyi kapsamak zorunda olma kısmını anlıyorum ama çocuğu sigara içen bir ebeveyn bu uygulamayı kullanır mı? Soruların biraz rafine edilmesi lazım. ”

- The comments related to medical content in the article 5:

Dentist 3: “Köpek ve arka azı dişlerde arka basamak, what? Ortodontik sorunu olan çocukları ebeveynleri çapraşıklık var diye getirirler bence, hekimler için ok ama ebeveynler için çok fazla.”

The comments on UX/UI:

- The UX/UI related comments in the article 1.a:

Dentist 2: “Başlangıç sayfası gayet sakin.”

Designer 1: “Profil oluştur diyoruz ya, bu çalışır muhtemelen. Ama benim bu uygulamayı yükleme amacımdan farklı bir “call to action”. Ben çocuğumun ağız sağlığı ile ilgili değerlendirme yapmaya geldim. Önce profil oluşturmam gerektiğini bu uzun yazıyı okuduktan sonra anlıyorum. Profil oluştur, değerlendirmeye başla gibi bir sonraki basamağı görebileceğim bir veri bana verilse daha iyi olur.”

Designer 2: “Profil oluştur butonu var ya, bold olması aslında hoş. Medikal olması hissini veriyor. Bir yandan da yukarıda bir ikon olsa, kaç tane profil ekleyebiliyorum onu görmek iyi olabilirdi. Netflix'te ekstra profil vardır ya boş slotları görmek iyi olabilirdi belki."

- The UX/UI related comments in the article 1.b:

Designer 1: "İsmi, doğum tarihi diyoruz. Cinsiyeti de diyebiliriz."

Designer 3: "Kız ve erkek gözümden kaçtı. Direk isme odaklandım. Acaba o cinsiyet sekmeleri pozisyon olarak orada mı olmalı diye düşündüm."

Designer 1: "Bir de profil oluşturuyoruz ya, ben buraya kimin ismini kimin fotoğrafını gireceğim? Kendimin olmayacağına dair tek işaret burada kız yazması. Çocuğunuzun profili gibi bir ifade koymak gerekir."

Designer 1: "Çocuğunuz kaç kilo doğdu sorusu ve bir önceki eleman arasındaki mesafe biraz daha fazla olmalı. Bir soru bitti diğer soru başlıyor hissini vermek için. Burada soruları kartlarla ayırmak mantıklı olabilir."

- The UX/UI related comments in the article 1.c:

Parent 1: "Ben onu Likert Scale sandım."

Parent 2: "Çocuğunuzun sağlığını nasıl buluyorsunuz diye sordu zannettim. 10 üzerinden değerlendirme yaptım…Sistemi şimdi anladım. Mesela numaraların altında noktalar cevapladığım sorularmış."

Dentist 3: "Likert ölçeği olduğunu düşünmüştüm."

Designer 1: "Swipe ile sağ sola giden bir şey ise bu rakamların üstte değil altta olması lazım."

Designer 3: "Aslında soru adımlarını görmek güzel ama öylesine konulmuş gibi."

- The UX/UI related comments in the article 1.e:

Dentist 3:"Hastalıkların listelenmesi baya iyi çünkü insanın hemen aklına gelmiyor. Kronik hastalıkları normal hayatın bir parçası olarak algılıyorlar ya. Ben mesela anemnez alırken, kalp şeker tansiyon var mı diye soruyorum hastaya, yok diyor, sonra kan sulandırıcı aspirin türevi bir ilaç kullanıyor musunuz diye sorduğumda evet diyor. Neden diyorum? Stent takılmıştı diyor. Kalp hastası olduğundan adamın haberi yok."

Designer 1: "Hastalıkların hepsini göstermek gerekmeyebilir. Ya da liste olarak gör gibi bir ibare ile drop down menü yapmak daha mantıklı. Onun için de ayrı bir sayfaya gitmek daha iyi olur bu arada."

- The UX/UI related comments in the article 2.a:

Dentist 1: "Sık kullanılanlar kısmını okuyunca ben hikayedeki semptomları buradakilere uydurmaya çalıştığımı fark ettim. Mesela hassasiyette de saydamlaşma denilen semptomda da buradaki renk değişikliğini işaretlemeye yeltendim. Belki search box'ı yukarı almak mantıklı olabilir."

Designer 2: "Sık karşılaşılan semptomları ilk seferde görüyorum. Buradan bir semptom seçiyorum. Sanki bu belirtiyle ilgili bir şeyler seçmem gerekiyormuş gibi geldi. Aşağı indim seçtiğim şey sayfadan kayboldu. Sonra alttaki listeden bir şeyler seçmem gerekti. Burada koptum. Nerede olduğumu anlamadım dememin sebebi buydu. Yukarıdaki listede başka bir şey seçseydim sanki aşağıdaki listede başka şeyler çıkacaktı gibi bir algı oluşturdu."

- The UX/UI related comments in the article 2.b:

Dentist 3: "Valla hiç dikkatimi çekmedi. Orada öyle duruyormuş gibi geldi üstüne basınca bunu bana açıklayacağı hissini vermedi."

Me: "İ işareti olsa info gibi o zaman ilgini çeker miydi? Ya da şöyle sorayım iç bükey neydi diye sorduğunda bunun cevabını şuradan alırım diyeceğin ne olurdu?"

Dentist 3: "Belki açıklama verebilirsin ama tabi bu benim deneyimim, belki ben basmıyorumdur… Benim üstüme bas demiyor."

Designer 3: "Burada fotoğraf mı istiyor benden?"

Me: "O ikon sana fotoğraf yükleme izlenimi mi verdi?"

Designer 3: "Vermedi ama doğru bir izlenim de vermedi."

Dentist 1: "Resim ikonunun belki renkli bir şey olması belki basmaya teşvik eder."

- The UX/UI related comments in the article 2.c:

Dentist 3: "Bu neden böyle oldu? X'ler çıktı."

Designer 3: "Bunlar neden çarpı oldu? Tahmin etmişim gibi hissettiriyor bana. Enteresan bir deneyim yaratıyor. Bu bildiri doğru değil bence. Çarpı olmamalı en azından."

- The UX/UI related comments in the article 3.a:

Parent 2: "Aslında yanında tik vardı ama bir iki defa ekleyemedim ya ondan ekleye basma ihtiyacı duydum. Önce ekleyi görüyorum ve ileriyi sonra görüyorum sanırım bir de ondan. Belki ikisi aynı hizaya gelirse basmayabilirdim."

Parent 3: "Şuan ileri basınca doğru seçmiş oluyor muyum?"

Dentist 1: "Onu seçtim bir daha eklemeye ihtiyaç var mı? Yok sanırım."

Designer 1: "Çocuğunuzun bulaşıcı bir hastalığı var mı? Hayır. Hayır işaretlenmeden geçti. Neyi işaretlediğimi anlayamadım."

Designer 2: "Ekle butonu bana eklemek için ona basmam gerektiği hissini oluşturdu. Sanki sepete ekle gibi. Tıklamazsam sanki kaydolmayacak gibi geldi."

Designer 3: “Seçtiğim şey nerede? Duruyor mu?... İlaçları eklediğimde bir yerde gözükse, ne eklediğimi bilsem.”

- The UX/UI related comments in the article 3.b:

Dentist 2: “Ana sayfada çok fazla tuş var bu beni biraz korkuttu. Benim amacım değerlendirme yapmak, şimdi hangisi oral muayene rehberi mi acil durum rehberi mi hangisi? Yoksa dental kayıt mı yapacağım?”

Designer 1: “Çocuğunuzun sistemik bir hastalığı var mı? Ben şu an kayboldum. Eğer ileri diyorsam bir ilerleme görmem lazım. Bu e-ticaret sitelerindeki gibi. Profil oluşturma %17 çok bir şey ifade etmedi. Burada bir navigasyon barı olmalı sanki…Listenin altına nasıl ineceğim? Bütün listeyi scroll down yapmak gerekiyor.”

Designer 3: “Şimdi beslenme alışkanlıklarına mı geçtim? Yeni stepler gördüm. Bu biraz kafa karıştırıcı olabilir. Hiyerarşisi tam anlaşılmıyor…Neredeydim ne yapıyordum unuttum gibi. Çocuğun profilini oluşturuyordum. Şimdi ne yapıyorum?”

- The UX/UI related comments in the article 4.a:

Parent 3: “Sonuçların olasılığı var mı?”

Dentist 1: “Bu sonuçlar risk sırasına göre mi bunların hiyerarşisi aynı mı? Belki renk kullanılabilir.”

Designer 1: “Burada drop down değil de, read more olmalı gibi geldi. Birkaç satır okumalıyım ve devamını okumak istersem basmalıyım.”

Designer 2: “Yüksek ve düşük olasılık bilgisi vermenin bir anlamı var mı? Bu şekilde bölmek sonuçları ve akordeon şeklinde açmaya gerek olmadığını düşünüyorum. Tek seferde okuyabileceğim bir sonuç sayfasını tercih ederim. Yüksek dediğin zaman ilk tıklayacağım şey o ve diğer seçeneklerden kopuyorum. Beni diğerlerini göz ardı etmememe sebep olmalı.”

Designer 3: “Sonuç sayfasında box var tıklıyorum, başka bir box açılıyor. Aslında bu mantıklı ama bunun üzerine tıklayacağım hissi vermiyor. Renkten dolayı olabilir. Belki call to action bir şey olabilir. Daha fazla bilgi al gibi olabilir. Read more vardır ya hani altı çizili. Böyle bişey olabilir. Neden alttaki boxlar outline da bu içi dolu?”

- The UX/UI related comments in the article 4.b:

Parent 1: “Sonuçlardan çok etkilendim. Bence baya iyi buldu. Hepsine tek tek bakar mıydım bilmiyorum ama yüksek olasılıklı olanlara bakardım. Bir de diş hekimine yönlendirmesi çok güzel. Bazen doktora gidersin, gidersin sorunun çözülmez. Belki de buradan araştırıp düşük olasılıklı sorunların üzerinde durabilirsin. O çok hoşuma gitti. Ben olsam bunu telefonumda tutarım. İhtiyacım olduğunda indirip sonra sildiğim bir uygulama olmaz.”

Parent 2: “Genel olarak iyi bu arada. Telefonda kullansaydım daha rahat olurdum. En baştaki soruyu anlamamıştım. Ama diğer tarafları bence kolay. Bazı sorular spesifikleşiyorsa o zaman çok daha iyi olur. Fotoğrafın olması çok avantajlı. İnsanlar gördüğünü daha iyi anlar. Ben olsam fotoğrafa bakarak değerlendirme yaparım. Onun dışında baya rahat aslında.”

Parent 3: “Açıkçası kullanılabilir bu ürün. İnsan kendi özelliğine göre referans bilgi almayı bekliyor. Hastalık tespiti için de görsel ve spesifik bir şey ya, Google’dan bakamıyorsun illa doktora gitmek zorundaymış gibi hissediyorsun. Bu açıdan değerli. Bir de bilgileri kaydetmek ve takip için aslında mantıklı. Rutin kontrol gibi şeyler de konulabilir. Benim çocuğum ufak olduğu şimdiye kadar ihtiyacım olmadı.”

Dentist 1: “Sonuçlarda çok fazla şey çıkıyor. Burada mükemmel doktoru sanki sisteme koymuşuz, bir sürü sonuç çıkarıyor ama sonucun ne olacağıyla ilgili muhakemeyi de kullanıcıya bırakmışız gibi. Sanki bu kullanıcıya haksızlık olurmuş gibi geldi. Sanki bu sonuçlar hekimlere gösterilse daha iyi olur gibi...Evet hukuksal sorumluluklar konusunda haklısın bir de diyelim ki bir sonuç çıkardı. Erozyon hakkında bilgi aldı, hekime gitti. Hekim bakıp çürük var dedi. O zaman ne olacak hayır bu erozyon mu diyecek? Burada birkaç olasılık olması mantıklı.”

Dentist 3: “Olası riskler sayfasında çok şey görmek ebeveynleri tedirgin edebilir, özellikle Türkiye’de.”

Designer 1: “Bana o kadar çok şey söyledi ki, sanki hiçbir şey söylemedi gibi geliyor… Bana süreç çok uzun geldi. Tabi doğru tanıyı koyması için bu bilgileri istemesini anlıyorum ama süreç yine de uzun ve sonunda bana diş hekimine gitmemi öneriyor. Ben zaten gitmemek için buraya başvuruyorum.”

Designer 2: “Ne yapmalıyım’a tıklıyorum.” (Gülüyor.) “Bu cevap benim en gıcık olduğum şey. Ne yapmalıyım? Diş hekimine başvurmalısın. Bana bir şey yapabilecekmişim gibi bir his uyandırıp beni sonunda zaten yapmam gerekene, bildiğim bir şeye yönlendiriyor…Ne yapmalıyım yerine başka bir şey yazılabilir.”

Designer 3: “Arayüzlerin draft halinde olduğu hissediliyor. Ama içerik yetkin bir şeyle karşılaştığını hissettiriyor.”

Appendix– B: Ethical Committee Approval

İTÜ

SAĞLIK VE MÜHENDİSLİK BİLİMLERİ İNSAN ARAŞTIRMALARI ETİK KURULU
İTÜ-SM.İNAREK

Onay No: İTÜ-SM.İNAREK-2022-10

24.11.2022

Yürütücülüğünü İstanbul Teknik Üniversitesi (İTÜ) Lisansüstü Eğitim Enstitüsü, Endüstriyel Tasarım Yüksek Lisans Programı öğrencisi Deniz Göçhan'ın yaptığı *'Bir Ağız Sağlığı Mobil Uygulamasının Geliştirme Sürecinde Kullanıcı Araştırmasının Rolü: Uzman Olmayan Kullanıcılar için Uzman Sistem Tasarlamak'* başlıklı proje başvuru dosyaları değerlendirilmiş ve etik açıdan uygun bulunmuştur.

Prof.Dr. Pervin BAŞARAN AKOCAK
İTÜ – Gıda Mühendisliği Bölümü

[redacted]

Mazereti nedeniyle katılmadı
Prof.Dr. Sibel Aylin UĞUR İŞERİ
İstanbul Ü. – Aziz Sancar Deneysel Tıp Araş Ens.

Doç.Dr. Ayça SAYI YAZGAN
İTÜ – Moleküler Biyoloji ve Genetik Bölümü

Dr. Öğr. Üyesi Alaeddin Burak İREZ
İTÜ – Makina Mühendisliği Bölümü

Mazereti nedeniyle katılmadı
Doç. Dr. Tuba YILMAZ ABDOLSAHEB
İTÜ – Elektronik ve Haberleşme Müh. Bölümü

CamScanner ile tarandı

CURRICULUM VITAE

Name / Surname : Deniz GÖÇHAN

EDUCATION :

- D.D.S. : 2008, Ege University, Dentistry Faculty
- Ph.D. : 2013, Baskent University, Institute of Health Sciences, Periodontology Department

PROFESSIONAL EXPERIENCE AND REWARDS:

- 2013 – 2016, Istanbul Aydin University, Dentaydin Dental Hospital, Oral Surgeon
- 2016 – 2018, Medical Research and Technology, Entrepreneur
- 2014 – 2020, Uludent Dental Policlinic, Oral Surgeon
- 2020 – Present, MYMOUTH Digital Health Technologies LTD., Founder

PUBLICATIONS, PRESENTATIONS AND PATENTS:

Göçhan, D., Oduncuoğlu, B. F., & Alaaddinoğlu, E. E. (2010). Two cases of Mucous Membrane Pemphigoid. *Turkish Society of Periodontology 40th Scientific Congress* (40th ed.). Turkish Society of Periodontology.

Atasoy Şentürk, R., Oduncuoğlu, B. F., & Göçhan, D. (2011). Platelet-rich fibrin applications: Case reports and short-term results. *Turkish Society of Periodontology 41st Scientific Congress* (41st ed.). Turkish Society of Periodontology.

Göçhan, D., Atasoy Şentürk, R., & Oduncuoğlu, B. F. (2011). Effect of low-level laser therapy on Oral Lichen Planus: Two case reports. *Turkish Society of Periodontology 41st Scientific Congress* (41st ed.). Turkish Society of Periodontology.

Göçhan, D., Oduncuoğlu, B. F., & Atasoy Şentürk, R. (2011). Free gingival graft harvested from hyperplasia on the edentulous alveolar ridge: Case report. *Turkish Society of Periodontology 41st Scientific Congress* (41st ed.). Turkish Society of Periodontology.

Göçhan, D., Oduncuoğlu, B. F., & Öncü, E. (2011). Osler Weber Rendu Syndrome: Case report. *Turkish Society of Periodontology 41st Scientific Congress* (41st ed.). Turkish Society of Periodontology.

Öncü, E., Alaaddinoğlu, E. E., & Göçhan, D. (2011). Different plastic surgical techniques for the treatment of multiple adjacent gingival recessions: Case report. *Turkish Society of Periodontology 41st Scientific Congress* (41st ed.). Turkish Society of Periodontology.

Özbek, N., Göçhan, D., Alaaddinoğlu, E. E., & Bulut, Ş. (2011). Effectiveness of chitosan film on free gingival graft donor site healing. *Turkish Society of Periodontology 41st Scientific Congress* (41st ed.). Turkish Society of Periodontology.

Hamzaçebi, B., Oduncuoğlu, B. F., Alaaddinoğlu, E. E., & Göçhan, D. (2012). Treatment of peri-implant bone defects with platelet-rich fibrin. *Europerio 7 - Clinical Research: Implant therapy outcomes* (7th ed.). European Federation of Periodontology.

Ünal, S., Oduncuoğlu, B. F., Öncü, E., Göçhan, D., Kutsal, D., Şahin, D., . . . Bulut, Ş. (2012). Perceptions of patients' own smiles, beautiful or healthy? *Europerio 7 - Clinical Research: Epidemiology and delivery of care* (7th ed.). European Federation of Periodontology.

Göçhan, D. (2013). Evaluation of gingival vascular structures in periodontal diseases via video-capillaroscopy technique (PhD Dissertation). Başkent University.

Göçhan, D., & Bağlı, H. H. (2017). Design thinking approach for a medical entreprise: A case about a point of care testing device. *Cross(-)abilities in Design* (14th ed.). 4T Design and Design History Society.

Göçhan, D., & Oduncuoğlu, B. F. (2019). Capillaroscopic evaluation of smoking and non-smoking patients' gingival margin. *Süleyman Demirel Üniversitesi Sağlık Bilimleri Dergisi, 10*(2), 105–110. https://doi.org/10.22312/sdusbed.521171