



Titre: Investigating User Perceptions and User Trust of Smart Health Applications
Title:

Auteur: Banafsheh Mohajeri
Author:

Date: 2023

Type: Mémoire ou thèse / Dissertation or Thesis

Référence: Mohajeri, B. (2023). Investigating User Perceptions and User Trust of Smart Health Applications [Mémoire de maîtrise, Polytechnique Montréal]. PolyPublie.
Citation: <https://publications.polymtl.ca/54868/>

 **Document en libre accès dans PolyPublie**
Open Access document in PolyPublie

URL de PolyPublie: <https://publications.polymtl.ca/54868/>
PolyPublie URL:

Directeurs de recherche: Jinghui Cheng
Advisors:

Programme: Génie informatique
Program:

POLYTECHNIQUE MONTRÉAL

affiliée à l'Université de Montréal

Investigating user perceptions and user trust of smart health applications

BANAFSHEH MOHAJERI

Département de génie informatique et génie logiciel

Mémoire présenté en vue de l'obtention du diplôme de *Maîtrise ès sciences appliquées*

Génie informatique

Août 2023

POLYTECHNIQUE MONTRÉAL

affiliée à l'Université de Montréal

Ce mémoire intitulé :

Investigating user perceptions and user trust of smart health applications

présenté par **Banafsheh MOHAJERI**

en vue de l'obtention du diplôme de *Maîtrise ès sciences appliquées*

a été dûment accepté par le jury d'examen constitué de :

Zohreh SHARAFI, présidente

Jinghui CHENG, membre et directeur de recherche

Heng LI, membre

DEDICATION

*To my incredible Mom
for her unwavering love, support, and encouragement*

ACKNOWLEDGEMENTS

With immense gratitude, I acknowledge the numerous individuals who have played a vital role in my journey at Polytechnique Montreal. First and foremost, I extend my heartfelt appreciation to my research director, Prof. Jinghui Cheng, whose unwavering support and invaluable guidance were instrumental in completing this study. His kindness and readiness to address my inquiries were truly invaluable, and I am deeply thankful for the opportunity to be part of HCDLab. Additionally, my warmest thanks go out to the HCDLab members, whose camaraderie and collaboration enriched my experience.

Next, I am blessed to have a loving and supportive family by my side, always there to lend a helping hand during challenging times. Your unwavering support throughout these two years will forever remain etched in my heart. I am also thankful to my dear friends, who have been a constant source of encouragement and laughter, making this journey memorable and enjoyable.

Finally, I extend my heartfelt gratitude to the esteemed jury members for their meticulous review of this thesis.

RÉSUMÉ

Les applications de santé intelligentes sont de plus en plus répandues aujourd’hui, offrant aux gens les moyens d’atteindre un large éventail d’objectifs de santé, y compris le suivi de la condition physique, le contrôle du régime alimentaire et le diagnostic préliminaire. Cependant, malgré leur popularité croissante, les défis auxquels sont confrontés les utilisateurs quotidiens d’applications de santé intelligentes ne sont pas suffisamment bien compris, ce qui entrave leur développement. Cependant, malgré leur popularité croissante, les défis auxquels sont confrontés les utilisateurs quotidiens d’applications de santé intelligentes manquent de compréhension globale, ce qui entrave leur plein potentiel. Cette thèse se penche sur cette question et se concentre sur l’examen des perceptions des utilisateurs des applications de santé intelligentes actuelles, en mettant particulièrement l’accent sur le facteur crucial de la confiance de l’utilisateur. confiance des utilisateurs. En analysant les commentaires des utilisateurs et en s’inspirant de la littérature existante, cette étude vise à explorer les orientations potentielles de conception qui peuvent renforcer la confiance et garantir des applications de santé intelligentes centrées sur l’utilisateur et dignes de confiance, centrée sur l’utilisateur et digne de confiance.

Dans la première partie de cette thèse, nous étudions les problèmes communs des applications de santé intelligentes les plus populaires en analysant les commentaires des utilisateurs sur les sites web de santé intelligente. en analysant les commentaires des utilisateurs sur le Play Store relatifs à la conception de l’expérience utilisateur. Cette exploration intégrée des applications de santé intelligentes les plus populaires vise à nous aider à découvrir les préoccupations communes du point de vue des utilisateurs et à inspirer la création de lignes directrices de conception pour améliorer l’expérience des utilisateurs d’applications de santé intelligentes. L’analyse des résultats a révélé que les opinions des utilisateurs sur les composants intelligents des applications mobiles de santé s’articulent autour de plusieurs catégories clés. Les utilisateurs apprécient la précision, la personnalisation et la facilité de saisie des données dans les applications de santé intelligentes. Ils apprécient les applications qui fournissent des fonctions fiables liées à la santé, permettent la personnalisation et sont faciles à utiliser. Les utilisateurs font également confiance aux applications qui font preuve de transparence, de fiabilité scientifique et de réactivité. En outre, nous avons découvert que le manque de confiance était un problème important rencontré par les utilisateurs lors de l’utilisation d’applications de santé intelligentes.

La deuxième partie de cette thèse explore plus avant la question de la confiance des utilisateurs.

teurs. Nous avons adopté le cadre de Toreini et al. [1] sur la fiabilité des systèmes basés sur l’IA et mené une recherche *speculative design* dans le contexte des applications de santé intelligentes. En particulier, nous avons développé quatre sociétés fictives qui reflètent différentes caractéristiques de ce cadre, ainsi que les prototypes d’applications de santé correspondants pour chaque société. Nous avons ensuite mené une étude auprès des utilisateurs afin d’évaluer ces sociétés fictives et leurs prototypes correspondants en termes de fiabilité et nous avons analysé les résultats. Les résultats confirment que l’équité, l’explicabilité, l’audibilité et la sécurité sont essentielles pour que les applications de santé intelligentes soient dignes de confiance. Les participants perçoivent les sociétés présentant ces caractéristiques comme moralement supérieures et plus dignes de confiance. En outre, la personnalisation et l’explicabilité sont apparues comme des facteurs essentiels, et l’autonomie et la liberté de l’utilisateur ont été fortement appréciées.

Dans l’ensemble, cette thèse étudie les perceptions des utilisateurs des applications de santé intelligentes, en se concentrant particulièrement sur la confiance des utilisateurs et l’amélioration de l’expérience utilisateur. Dans la première phase, nous avons analysé les commentaires des utilisateurs, en identifiant les préoccupations concernant l’exactitude, la personnalisation et la saisie des données. Dans la deuxième phase, nous avons utilisé la conception spéculative pour explorer l’importance de l’équité, de l’explicabilité, de l’audibilité et de la sécurité dans l’établissement de la confiance. Cette thèse contribue au domaine en fournissant des lignes directrices pour la conception d’applications de santé intelligentes dignes de confiance, en mettant l’accent sur la personnalisation et l’explicabilité, et en soulignant l’importance de l’agence de l’utilisateur et des intentions perçues dans la promotion de la confiance.

ABSTRACT

Smart health applications are increasingly prevalent today, offering people the means to achieve a wide range of health goals, including fitness tracking, diet monitoring, and preliminary diagnosis. However, despite their growing popularity, there is a lack of comprehensive understanding of the challenges faced by everyday users of smart health apps, hindering their full potential. This thesis delves deep into this issue and focuses on examining user perceptions of current smart health apps, with a particular emphasis on the crucial factor of user trust. By analyzing user reviews and drawing insights from existing literature, this study aims to explore potential design directions that can enhance trust and ensure user-centric, trustworthy smart health applications.

In the first part of this thesis, we investigate the common issues of popular smart health applications by analyzing users' reviews on the play store related to user experience design. This integrated exploration of popular smart health applications aims to help us discover the common concerns from the users' perspective and inspire the creation of design guidelines to improve the user experience of smart health apps. The analysis of the results revealed that users' opinions about the smart components of health mobile apps revolved around several key categories. Users valued accuracy, customizability, and convenience of data input in smart health apps. They appreciated applications that provided reliable health-related functions, allowed personalization, and were easy to use. Users also trusted apps that demonstrated transparency, scientific reliability, and timely responses. Additionally, we discovered a lack of trust as an important issue encountered by users when using smart health applications.

The second part of this thesis further explores the issue of user trust. We adopted Toreini et al.' [1]'s framework of the trustworthiness of AI-based systems and conducted a *speculative design* research in the context of smart health apps. Particularly, we developed four fictional societies that reflect different features of this framework, as well as the corresponding health app prototypes for each society. We then conducted a user study using to evaluate these fictional societies and their corresponding prototypes in terms of trustworthiness and analyzed the results. The results confirm that fairness, explainability, audibility, and safety are crucial for trustworthy smart health applications. Participants perceived societies with these characteristics as morally superior and more trustworthy. Additionally, personalization and explainability emerged as vital factors, and user agency and freedom were highly valued.

Overall, this thesis investigates user perceptions of smart health applications, particularly focusing on user trust and improving user experience. in the first phase, we analyzed user

reviews, identifying concerns about accuracy, customizability, and data input. In the second phase, we employed speculative design to explore the importance of fairness, explainability, audibility, and safety in establishing trust. This thesis contributes to the domain by providing design guidelines for trustworthy smart health apps, emphasizing personalizability and explainability, and highlighting the significance of user agency and perceived intentions in fostering trust.

TABLE OF CONTENTS

DEDICATION	iii
ACKNOWLEDGEMENTS	iv
RÉSUMÉ	v
ABSTRACT	vii
TABLE OF CONTENTS	ix
LIST OF TABLES	xii
LIST OF FIGURES	xiii
LIST OF SYMBOLS AND ACRONYMS	xiv
LIST OF APPENDICES	xv
CHAPTER 1 INTRODUCTION	1
1.1 Context and Motivation	1
1.2 Objectives and Approach	2
1.3 Thesis Plan	2
CHAPTER 2 LITERATURE REVIEW	4
2.1 Previous Literature Related To The First Stage of The Project	4
2.1.1 Mobile Health Apps	4
2.1.2 Design with Smart Components	4
2.1.3 App Review Analysis	5
2.2 Previous Literature Related To The Second Stage of The Project	5
2.2.1 Trust Formation	5
2.2.2 Trust In Smart Systems	6
2.2.3 Trustworthy Health Apps	7
2.2.4 Speculative Design	7
CHAPTER 3 FIRST STUDY: INVESTIGATING THE COMMON USERS' ISSUES WITH POPULAR SMART HEALTH APPLICATIONS	10

3.1	Introduction	10
3.2	Methodology	11
3.3	Results	13
3.3.1	RQ1: Users' opinions about the smart component of health mobile apps	13
3.3.2	RQ2: Other factors regarding the design of smart health mobile apps in user reviews	16
3.4	Discussion	18
3.5	Conclusion	19
CHAPTER 4 SECOND STUDY: INVESTIGATING THE ESSENTIAL CRITERIA OF TRUSTWORTHY HEALTH APPLICATIONS BY APPLYING SPECULA- TIVE DESIGN		20
4.1	Introduction	20
4.2	Iterative Speculative Design of the Prototypes	22
4.2.1	Society Design	22
4.2.2	Initial Prototype Design And Design Iterations	26
4.3	Final Prototype Design	28
4.4	User Study Methods	36
4.5	User Study Results	39
4.5.1	Previous Experience With Smart Health Apps	39
4.5.2	Evaluation Of The Society Descriptions	40
4.5.3	Evaluation Of The Prototypes	44
4.5.4	Match Between Society Descriptions And Prototypes	47
4.6	Discussion	48
4.6.1	The Importance Of The Four Criteria	48
4.6.2	The Vitality Of Personalizability And Explainability	49
4.6.3	Giving Users Freedom	50
4.6.4	Design Considerations For The Development Of Trustworthy Applica- tions	51
4.7	Conclusion	52
CHAPTER 5 CONCLUSION		54
5.1	Summary of the Work	54
5.2	Limitations and Future Research	55
REFERENCES		57

APPENDICES	62
----------------------	----

LIST OF TABLES

Table 4.1	Profile of the participants of our user study	38
-----------	---	----

LIST OF FIGURES

Figure 3.1	The frequency of the categories related to the smart component . . .	14
Figure 3.2	The frequency of the other categories mentioned by the users	17
Figure 4.1	The old color theme of the prototype 1	29
Figure 4.2	Adding additional recipe options to prototype 3	29
Figure 4.3	Prototype 1	31
Figure 4.4	Prototype 2	33
Figure 4.5	Prototype 3	35
Figure 4.6	Prototype 4	37
Figure 4.7	The ranking of societies in terms of trustworthiness	45
Figure 4.8	The ranking of prototypes in terms of trustworthiness	48

LIST OF SYMBOLS AND ACRONYMS

UI	User Interface
UX	User Experience
AI	Artificial Intelligence
ML	Machine Learning

LIST OF APPENDICES

Appendix A	The Prototypes	62
Appendix B	The Test Script For The User Study	84
Appendix C	The Code Book	90

CHAPTER 1 INTRODUCTION

1.1 Context and Motivation

The market is flooded with a vast array of health and fitness appliances, many of which incorporate smart components [2] [3]. While significant efforts have been made to develop and refine individual apps catering to specific health and fitness needs [4], there is a dearth of literature examining the broader challenges faced by everyday users of smart health apps that prevent them to fulfill their potential.

Moreover, smart health apps, like other technologies, are undergoing rapid and unforeseeable evolutions. These advancements encompass the integration of new technologies and the addition of functionalities such as adaptive and personalized health [5], sophisticated artificial intelligence [6], gamification and behavioral incentives [7], integration with electronic health record systems [8], and the introduction of data analytics and predictive models [9]. These new features open up exciting possibilities and enable health applications to offer services that were previously unimaginable. However, this future landscape also presents new risks and challenges. Given this dynamic environment, it becomes crucial to explore how these future developments align with users' current needs and concerns.

This thesis aims to bridge the gap between evolving technologies and users' needs to create a more positive and beneficial landscape for smart health applications. Particularly, we first focused on identifying the challenges perceived by users of smart health applications across different health domains, by analyzing user reviews of commercially available smart health apps. Through this analysis, we uncovered an important area of concern for users: the issue of **trust**. Building upon this initial study and drawing insights from existing literature, we then aim to examine the potential directions in which smart health applications may evolve in the future and the impact of these directions on user trust. This second study aims to identify the most suitable directions for developing trustworthy smart health applications while exploring the criteria that make them compatible with the users' needs and expectations. Recognizing that the development of smart health applications, like other technological advancements, is deeply influenced by broader societal contexts and the perspectives of various stakeholders involved in their creation and use [10], we employ the innovative method of *speculative design* to envision alternative societal and technological paths. Speculative design, as defined by Dunne and Raby [11], is an approach to design that explores and imagines alternative futures through the creation of provocative and thought-provoking design artifacts. It aims to challenge existing assumptions, question societal norms, and stimulate critical discourse

about the implications of emerging technologies and possible future scenarios. The ultimate goal of our research is to provide insights into the design of user-centered and trustworthy smart health applications that effectively address the evolving needs of individuals in this dynamic domain.

1.2 Objectives and Approach

The objective of this thesis is to investigate the common concerns of users of smart health applications and explore the potential directions for the development of trustworthy smart health applications. The study aims to identify the main user concerns that contribute to the user experience of these applications. Additionally, the study aims to assess the applicability of the criteria of *fairness*, *explainability*, *audibility*, and *safety*, as indicated in Toreini’s framework [1], in the context of smart health applications and explore the possibility of additional elements that contribute to the design of trustworthy smart health applications.

- *User Review Analysis*: The study began by analyzing user reviews of popular health applications with smart components. This analysis identified the common concerns expressed by users regarding these applications. The focus is on understanding the main concerns and issues of smart health apps from the users’ perspective.
- *Speculative Design and User Study*: Building upon the insights gained from the review analysis, this study adopted a speculative design approach to envisioning potential future trajectories for smart health applications concerning user trust. To achieve this, fictional societies were created, each exhibiting distinct variations in fairness, explainability, audibility, and safety criteria. These hypothetical societies served as a foundation for exploring the trustworthiness of smart health applications in diverse contexts. Subsequently, a series of prototypes, corresponding to each fictional society, were developed through an iterative process. To assess the prototypes and society descriptions, a set of user studies were conducted. These studies aimed to investigate the critical criteria contributing to the trustworthiness of smart health applications and to gain insight into users’ opinions and experiences regarding the trustworthiness of such applications.

1.3 Thesis Plan

The structure of the thesis is as follows. In Chapter 2, we review the previous literature related to the two stages of this thesis. In Chapter 3, we present the first stage of this

study, which investigates the common user concerns of popular smart health applications. In Chapter 4 and by building on the results of the first study and the importance of trust, we report on the second phase of this study exploring the essential elements of trustworthy smart health applications by applying a speculative design approach. Finally, Chapter 5 presents the overall conclusions and discusses the limitations and possible future research related to the entire study.

CHAPTER 2 LITERATURE REVIEW

2.1 Previous Literature Related To The First Stage of The Project

In this first phase of the study, our research has some common ground with works done regarding mobile health apps in general, design practices with smart components, and app review analysis. Below is a brief overview of each group:

2.1.1 Mobile Health Apps

With the continuous surge of new applications, there have been multiple studies done on specific types of applications serving a particular health need or group of users. For example, Niess et al. [12] conducted an interview-based study to understand attitudes and practices related to fitness-tracking apps in a global setting. They proposed design suggestions to help build more inclusive fitness-tracking apps. Lau et al. [13] conducted a systematic review of publicly available apps for psychosocial wellness and stress management. They found that only a very small fraction of such apps are backed up with scientific evidence of their effectiveness. Kao et al. [14] reviewed a wide range of mobile health apps and identified three major barriers to advancing the potential of mobile health apps, including (1) lack of regulatory oversight, (2) limited evidence-based literature, and (3) concerns of privacy and security. Despite these previous efforts, There has been little research done particularly on the overall landscape of user opinions on health applications with a smart component. This study aims to fill this gap.

2.1.2 Design with Smart Components

Many studies have been conducted to understand and address the new challenges when designing systems and applications that involve smart components. For example, Dove et al. [15] examined the design practice for applications with machine learning (ML) techniques by conducting a survey with UX practitioners. They questioned whether or not these practitioners took the ML component into account during their design process and discussed what new research and curriculum can help make the ML parts a more integral design material. Amershi et al. [16] leveraged the rich literature in the history of HCI research and proposed 18 general design guidelines concerning learnability, error recovery, and evolution of AI systems. Volkel et al. [17] aimed to decipher the essence of what makes a user interface intelligent by using a bottom-up approach to review the meaning of intelligence in user interface design

literature and proposed their own definition of intelligence in UI. Our study is informed by this body of literature to examine user concerns of smart health apps.

2.1.3 App Review Analysis

Previous studies have used app reviews as a proxy to understand users' concerns about various types of apps. Vasa et al. [18] conducted quantitative analysis on a large number of apps on the Apple App Store; they found that the reviews are generally short and that lower ratings were correlated with longer reviews. Ghosh et al. [19] qualitatively analyzed reviews written by children for online safety apps that are focused on parental control; they found that children considered these apps invasive to their privacy and damaging parent-children relationships. Viswanath et al. [20] developed a framework describing factors influencing user adoption of mental health conversational agents through a qualitative analysis of user reviews of such systems. Phillips et al. [21] used user reviews of commercially available video games to identify games that have potential therapeutic values. Our methods are inspired by these previous studies.

2.2 Previous Literature Related To The Second Stage of The Project

Our work in the second phase of the study is most closely related to previous studies that focused on (1) trust formation, (2) trust in smart systems, (3) trustworthy health applications, and (4) speculative design. We briefly review each group of literature in the following sections.

2.2.1 Trust Formation

Simpson [22] proposed that trust develops when the trusting partner comes to believe in the trustee after their initial act of trust is met with a trust-affirming action from the trustee. This crucial moment occurs when the trustee, facing a situation where they could choose to fulfill or betray the trust, opts for trustworthiness, thereby rewarding the initial act of trust. Through this process of positive reinforcement, the trusting partner gradually builds confidence in the trustee, leading to the establishment of trust in their relationship. The initial decision to trust a person or object could be the result of the conscious or unconscious signals sent by the trustee during the initial phase as well as the trustor's personal and cultural character. In short, trust is defined as a process of uncertainty reduction as individuals move from having confidence in their partner's general predictability to having confidence in their relationship values, motives, goals, and intentions. The same phenomena of trust formation after a risk-involved situation during which the trust can be either rewarded or violated

has also been mentioned by Ashraf et al [23]. Here, however, the authors mention that in addition to the initial calculated expectation of the trust being returned and fulfilled, there is an element of unconditional kindness and belief in the general altruism of people involved which can be a result of psychological benefits of such belief in group interaction.

2.2.2 Trust In Smart Systems

Siau et al. [24] believe that trust in technology is dependent on three factors: the individual factors regarding the user’s personality and ability to trust and take risks, the environmental factors such as the task at hand, and the situational and cultural credentials of the team or institute behind the technology and the technological factors such as its abilities and purpose. They also mentioned that trust in smart applications containing an AI element is usually built in two stages. The initial stage is where some easy-to-recognize factors such as the visual representation, the initial performance, the other users’ opinions, and the explainable nature of the system are important. The second and developmental stage is reliant on characteristics that are easier to discern through time such as long-term performance, reliability, safety, and privacy.

Tripp et al. [25] proposed that individuals tend to perceive technologies with greater social presence, social affordances, and affordances for social behavior as more human-like compared to technologies lacking these attributes. Moreover, for technologies perceived as having more humanness, human-like trusting beliefs, such as perceived usefulness, good intentions, and enjoyment from the interaction, hold greater significance than system-like trusting beliefs.

In this study, we adopt Toreini et al.’s framework [1] as it provides a robust foundation for defining the essential criteria of trustworthy smart technologies. They developed this framework by reviewing relevant literature and building upon it. The foundation of their approach lies in the incorporation of the *ABI+ framework* of trust, encompassing elements such as ability, benevolence, integrity, and predictability. Furthermore, they enriched this methodology by creating a recent mapping that links *ABI+* qualities to technologies that foster trust. Specifically, within the domain of machine learning, they identified four key categories of trustworthiness technologies: Fairness, Explainability, Auditability, and Safety. Subsequently, they conducted an analysis to examine how these essential qualities of technologies contribute to overall trustworthiness. A detailed explanation of these four criteria is presented in Chapter 4.

2.2.3 Trustworthy Health Apps

Some studies have focused on the regulatory and societal possibilities for creating trustworthy health apps, Wicks et al. [26] discussed that a framework consisting of the following steps of boosting app literacy, creating app safety consortium, enforcing transparency, having an active medical review and government regulation could lead to the development of trustworthy health apps. Similarly, Albrecht et al. [27] contended that mandating health apps to disclose a comprehensive document detailing their data processing and collection methods, the app’s intended purpose, commercial information, and other relevant data can foster user trust. By providing users with access to this information, individuals can make informed decisions, allowing them to observe and evaluate the app’s practices, thus increasing their confidence and trust in the app.

Some other works studied the trustworthiness of smart health apps in the context of a specific domain or function. As an example, Torous et al. [28] examined the current hurdles preventing mental health professionals and regular users from trusting mental health apps and found that many of the available health applications do not perform their promised functionality and can not be trusted with protecting the user’s privacy.

2.2.4 Speculative Design

Speculative design is an innovative approach that explores and imagines alternative futures by creating provocative and thought-provoking design artifacts. It challenges existing assumptions, questions societal norms, and stimulates critical discourse about the implications of emerging technologies and possible future scenarios [11]. The speculative design offers a creative space for designers and researchers to envision and explore possibilities that may not exist in the present reality, enabling them to gain insights and generate ideas that can inform future design decisions. While speculative design has gained prominence in research relatively recently, there are numerous exploratory and constructive studies that have employed this method. Auger et al. [29] provided a foundational understanding of speculative design by presenting several case studies that highlight the importance of engagement methods and the presentation manner in successfully creating and managing speculation in a useful and thought-provoking way.

In the realm of healthcare, speculative design has been applied to explore and address various health-related issues. For example, Hoang et al. [30] utilized speculative design in the form of a multi-choice storytelling tool to help patients create their own narratives when working with health tools. By enabling patients to share their experiences through story-

telling, this approach aimed to enhance understanding of patients' needs and facilitate better communication and collaboration in healthcare settings.

Ray et al. [31] explored the use of speculative design fiction to promote positive social behaviors during health and environmental concerns. Through speculative design, the authors created story structures and interactions that encouraged cooperation and societal good. They engaged in workshops to collaboratively develop storytelling approaches aimed at fostering sustainable behavioral patterns for positive social impact.

Smith et al. [32] investigated the role of speculative design in challenging and reimagining healthcare algorithmics. They discussed how data-based algorithmic prediction tools in healthcare could perpetuate existing normativity and exclusions. By adopting a posthuman speculative ethics approach, they used speculative design to envision alternative futures and challenge traditional notions of ethics and care in healthcare algorithmics. Their goal was to foster more inclusive and ethically conscious healthcare practices by critically engaging with predictive management through speculative design.

As another example of applying the speculative design approach to a specific health-related service, Hebrok et al. [33] discussed the use of design fiction to critically reflect on sustainable food consumption. They introduced the design fiction "Bird," a food delivery service that provides personalized food rations based on nutritional needs and self-improvement goals. Through this design fiction, the authors explore possible sustainable food futures and encourage reflection on ethical, social, cultural, political, and environmental aspects of food consumption.

Overall, speculative design offers a valuable tool for envisioning and exploring potential futures in multiple contexts including that of healthcare and smart health applications. [34] By going beyond the constraints of the present and engaging in imaginative thinking, researchers and designers can generate insights, challenge assumptions, and push the boundaries of health care and health-related services [35].

The existing literature discussed above provides substantial research on the subject of this study. However, there remains a noticeable gap in previous work, as no comprehensive investigation into the factors influencing the development of trustworthy health applications has been conducted. While some studies have examined the trustworthiness of specific groups of health apps, their narrow focus limits the overall scope and generalizability of the findings.

Moreover, the current research on smart health apps tends to be primarily centered around technical limitations, overlooking the broader design frameworks needed to adapt to the rapidly evolving landscape of smart technologies. Consequently, there is a lack of research

exploring the holistic approach required to create trustworthy health apps in this dynamic environment.

In light of these research gaps, this study aims to take an initial step toward addressing these limitations by exploring a more comprehensive perspective on the factors contributing to the trustworthiness of health applications. By delving beyond the technical aspects and considering the broader design frameworks, this research endeavors to provide valuable insights for fostering trust in health apps amidst the ever-changing landscape of smart technologies.

CHAPTER 3 FIRST STUDY: INVESTIGATING THE COMMON USERS' ISSUES WITH POPULAR SMART HEALTH APPLICATIONS

3.1 Introduction

Mobile health applications are increasingly popular today. They can help individuals with numerous types of health-related objectives such as weight loss, self-management of health problems, and self-diagnosis, to just name a few. With the growth of the health applications market, there are more and more applications integrating some form of smart components, leveraging techniques such as machine learning, in order to offer capabilities of progressively adapting to individual users' needs. However, preliminary research demonstrated that many users abandon health apps after the initial enthusiasm [36] in part because the users do not possess confidence in the applications' ability and they are not satisfied with the user experience that they receive. In particular, smart health applications, although show great potential, can present new and unique challenges to the user experience. These challenges originated partly from the algorithms that enable the smart capacity, whose performance is not always accurate and even predictable. If not communicated appropriately, this problem can discourage the users from continuously using the application as it erodes their trust.

On the other hand, a well-designed application conveying adequate information about the smart component, as well as its corresponding capabilities, can enhance the user experience. Sufficient user control and feedback can also help improve the abilities and performance of the smart component [16]. There have been previous studies aimed at (1) understanding user perceptions and opinions towards apps for a specific health purpose (e.g., [12]) and (2) reviewing the landscape of different kinds of commercially available health apps, particularly to the smart components that are increasingly prevalent in those apps. This knowledge is important in advancing the design, promoting the adoption, and thus fulfilling the potential and promises of smart health apps. So in this study, we aimed to take a step toward addressing this gap by analyzing the users' reviews of popular commercially available smart mobile health applications. Through this analysis, we harvest the overall concerns related to both positive and negative opinions about the design decisions and capabilities of the applications affecting the user experience, especially those related to the smart component. In particular, we target two main research questions:

- RQ1: What are the users' opinions about the design decisions related to the smart component of mobile health apps?

- RQ2: What other factors regarding the design of mobile health apps are discussed in the user reviews?

In order to answer these research questions, we conducted a qualitative analysis of the top most relevant reviews for each app. We found that understandably, the smart component plays a major role in the users’ impression of the applications. Users especially care about its accuracy and ease of use and will get particularly frustrated when the smart component is not accurate on many occasions and on these occasions without any explanation.

3.2 Methodology

To answer our research questions, we sampled popular health mobile applications that included a smart component and analyzed the user reviews of these applications. We intentionally chose to examine the most popular health applications to gain a comprehensive understanding of general user perceptions. These popular apps have the largest user base, allowing us to focus on issues that are most relevant to a broad audience. However, it’s important to acknowledge that this approach has its limitations. Specifically, focusing solely on popular apps may result in overlooking issues unique to less popular or specialized health applications. We focused on Android applications in the Google Play store because it is considered the most popular mobile platform [37]. We sampled applications according to the following steps. First, the applications are selected from the Health & Fitness and Medical categories of the Google Play store. Second, we collected the applications that had more than 10 million installs to ensure that the analyzed applications had a large user base; 6 apps were collected in this step. We then inspected the main functionalities of each application by installing them and using them for a few weeks, to filter out any applications that did not have any smart components.

We considered a component “smart” if it is capable of being personalized and adjustable to the user’s choices and/or usage history. Although we cannot be certain if this is achieved through machine learning or other techniques, the capability of learning and adjusting makes the application appear to be intelligent from the users’ perspective. Finally, to avoid homogeneous results we included only one application with a specific health function (e.g., diet coach, step tracking, etc.). The final selection of applications was strategically limited to maintain a manageable scope for our study. At the same time, we aimed to include at least one application from each key sub-category of smart health applications. These sub-categories encompass diet and workout planners, mental health and meditation platforms, diagnostic tools, and activity trackers. These steps resulted in the following six applications

that were the focus of further analysis:

- *Google Fit: Activity Tracking*: This app collects multiple data from sensors on Android devices, such as the number of steps, heart rate, sleep cycle, calorie consumption, etc. Its smart component calculates a “Heart Point” according to the gathered information based on the individual user’s activity. The higher the “Heart Point” number the more active and healthy the user had been according to the app. Users can also set “Heart Point” goals for a time period (daily, weekly, ...) and view their progress on the app.
- *Six Pack in 30 Days*: This app receives users’ health-related information like their weight, age, body fat, and fitness target to produce a personalized workout plan using the smart component. The workout plan can be further customized and the users can choose between multiple options. The app also keeps track of users’ related activities such as the amount of water or calorie consumed, the fluctuations of their weight, and their general fitness progress.
- *Headspace: Meditation & Sleep*: The app employs information about the users’ personal data and their schedule, as well as their mental struggles and goals to provide users with a personalized plan. The plan consists of different mental health practice sessions and guides on top of other useful tools like relaxing music and bedtime stories.
- *YAZIO Fasting & Food Tracker*: The app uses users’ personal data and their dietary targets to create customized meal plans. Additionally, the app has the ability to track the number of calories, carbohydrates, proteins, and fat consumed by the users through a smart scanner to recognize this information from the most common food products. It also has the ability to create personalized fasting schedules for users who are engaged in fasting.
- *WebMD: Symptom Checker*: This app provides users with a range of medical-related functions. It can receive a set of symptoms from users and predict possible diagnoses. Additionally, users can use the app to keep track of their medications, including the side effects and the possible interactions with other medications. The app also allows users to discover medical professionals and specialists available to them and make appointments.
- *Step Tracker - Pedometer*: The app uses built-in sensors of mobile devices to track the steps taken by the users. From this, it calculates information such as burned calories, walking distance, and exercise time. The app visualizes all of this information in a series of graphs. After selecting the applications, we collected the top 20 user reviews

of each application sorted in the most relevant mode of the Google Play store, resulting in 120 user reviews. We then conducted a thematic analysis of these user reviews

Much like our approach to selecting the number of applications, the number of reviews collected for each app was also chosen to balance time constraints with the need for meaningful data. We initially gathered 10 reviews for each application and incrementally added more until no significant new issues emerged in the additional reviews. This iterative method led us to a final sample size of 20 reviews for each application, allowing us to adequately capture the primary factors affecting user experience.

To answer our research questions. We employed a series of iterative open coding cycles on the user reviews, allowing for the application of multiple codes to each review. This iterative process continued until we reached a point where no further meaningful modifications to the codes or their groupings were necessary. Then in several extensive meetings, we merged and categorized the codes into higher-level themes with the help of an affinity diagram activity. These themes were then mapped to the research questions to provide specific answers.

Comments influenced by the smart component were initially categorized under a broad label called “Opinions about the smart component of health mobile apps.” These comments covered a wide range of topics. For instance, some addressed the accuracy and correctness of specific features, like inaccurate step counting or flawed path generation, both of which often used machine learning algorithms. Others focused on the smart component’s ability to deliver on an app’s advertised services, such as critiques of a diet plan’s quality in a diet app or the effectiveness of a workout schedule in a smart fitness coach app. Additional comments discussed the usability and aesthetic design, highlighting issues like unclear recommendations given to users or an unfriendly look of a smart coach within the app. After sifting through these comments, the remaining sections of the reviews, which still provided valuable but unrelated insights into users’ perceptions, were labeled as “Other factors regarding the design of smart health mobile apps.”

3.3 Results

3.3.1 RQ1: Users’ opinions about the smart component of health mobile apps

We found that users frequently mentioned smart components in their reviews. We grouped these concerns into the following categories.

(1) Accuracy. When the users felt that the application is accurate and can produce the correct information, they generally liked the app more and put more trust in it. For example, a Step

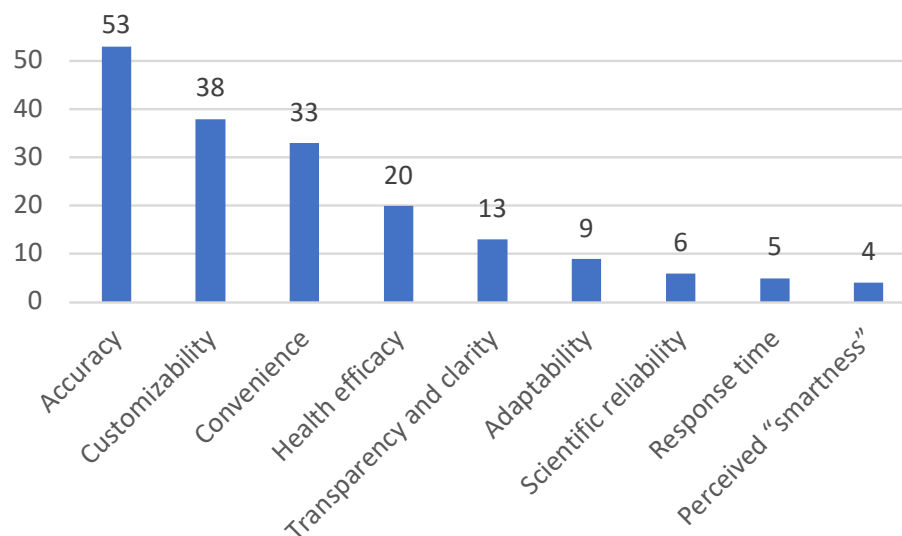


Figure 3.1 The frequency of the categories related to the smart component

Tracker - Pedometer reviewer noted that they like it because “Step counts, time recording, and calories burned are accurate.” On the flip side, one of the biggest complaints of the users was the lack of accuracy of the smart component. Sometimes the complaint was about how the app miscalculates the data that it receives from the users, such as their number of steps or the distance that they have walked. Other times the complaint was about how the app failed to record user activities completely. As an example of the latter case, a user reviewing GoogleFit mentioned: “It will count my steps when I’m standing still but not when I’m actually walking. I just lost 90 minutes of logging because it wouldn’t count my steps.” This inaccuracy also sometimes caused the users to lose their faith in the application and no longer be able to trust it. For example, another GoogleFit reviewer detailed this point of view by saying: “Unreliable and inaccurate tracking. Heart points disappear all of a sudden. On the day of the workout it shows the heart points, and the next day it becomes 0. Not recommended to track your daily life.”

(2) Customizability. The users appreciated it when the applications support users to make changes to the plan or the configuration of the smart component based on their personal preferences. They wanted not to be stuck with the default recommendation by the algorithms. They generally thought a more personalized plan suited more for their health needs and preferred to have control over the recommendations and suggestions. A Headspace user explained how the meditation sessions are particularly helpful for them in part because “being able to choose how long they last is great.” Predictably, users did not like applications that were not able to provide them with smart services based on their specific health-related

needs. Applications that failed in this regard sometimes would become useless for users as they were not able to achieve their personal health and fitness goals with them. For example, a YAZIO reviewer mentioned they could not really use the app since they do not “see an option to change how much you want to lose and how fast.”

(3) Convenience of data input. Many users put a great deal of emphasis on how easy it is for them to use the application without having to invest too much time or manual labor. The most prominent problem the users mentioned was that the applications were hard to use and took too much time to just be able to function, to the point that for some users they were not a really viable option anymore. Many occurrences of this problem happened with applications that required the users to fill in a very detailed part of their activities and habits such as every food intake or every exercise manually. For example, a YAZIO user noted how they are not very inclined to keep using the application because it is “Too cumbersome to actually track food every day”

(4) Health efficacy. Users complimented the applications whenever they had a smart component that was able to reliably perform the intended health-related functions. For example, the users of Headspace noted how it helps them “to get to sleep” and “concentrate when I need to” by providing good suggestions, sleep casts, and programs. Another user of Headspace corroborated the ability of the app to help users with addressing multiple mental health-related aspects effectively by saying: “Highly recommend for stress, anxiety, and focus! they also address pain, anger management, depression and so much more.”

(5) Transparency and clarity. The users generally tended to like the application more when they could figure out the required information and understand how the algorithm works without trouble and could operate the application hassle free. A user explained why they chose Step Tracker- Pedometer application by saying how “it has easy-to-read charts to help you visualize your progress”. On the other hand, users tended to complain when they found the application hard to understand, saw the user interface as unclear and confusing, or thought that they had to put a lot of work into understanding it.

(6) Adaptability. The users appreciated the applications that provided suggestions adapted to the user’s past information and behaviors. As an example, a Headspace user complimented the application by stating: “App is smooth to use and allows you to keep track of your stress, then suggests programs that could help at the moment”.

Similarly, users tend to like applications that can make changes to their programs or routine based on the feedback given by the users. A user reviewing the Six Pack in 30 Days application counted its ability to change based on the user’s feedback as a positive point by saying “The coaching system can adjust the workouts based on the feedback you give.”

(7) Scientific reliability. Users trusted applications more when they felt that the plans and services provided for them with the smart component have been built with consultation from health professionals and/or based on scientific data. These factors make the users feel that the app is safe and effective to use. A user reviewing Sixpack in 30 days highlighted this fact by mentioning that they trust the application because they feel like it gave them a “systematic and scientific way of reducing your fat and getting a toned body.”

(8) Response time. Users made complaints when the smart component took too long to respond and made the whole application slow. For example, a reviewer mentioned how they are uncertain about continuing to use YAZIO because for them it “feels slow”. (9) Perceived “smartness”. Users liked the applications that felt smart to them. Sometimes they mentioned that they can trust the app’s guidance like a real human professional. For example, a user noted how much Six Pack in 30 Days feels like a real coach by saying it feels like “You have your virtual trainer with you as if you are in a gymnasium”. Some users preferred that in case of uncertainty, the smart component showed them multiple options instead of just one; some were upset when it failed to do so. One user noted this problem when reviewing WebMD by complaining about how when it is not certain “it doesn’t give you any possible results.”

3.3.2 RQ2: Other factors regarding the design of smart health mobile apps in user reviews

Although less frequently, other factors that are not directly related to the smart component were also mentioned by the users in the reviews. Figure 1b shows the frequency of those factors in our sample.

(1) Monetary aspects. Users generally disliked apps with no free option or free trial. They were also sometimes irritated when there were too many ads. As an example, a user noted how they are discouraged to use WebMdD because the “Home page looks like an ad flier”. On the other hand, users were happy whenever they found an application with free options or a good free trial period. A user noted that they are satisfied with the free version of YAZIO calorie counter by saying “I didn’t pay for pro but the free is just as good.”

(2) Feature completeness. Many users expressed in the reviews that they desired to have certain features or felt that a feature was missing in the application. A common occurrence of this group stemmed from the users who wanted to connect the applications to another device like a smartwatch or another health application but were unable to do so.

(3) System failure. Users complain about many types of system failures, crashes or errors, the cause of which was unknown to the users. These issues were really bothersome for users since

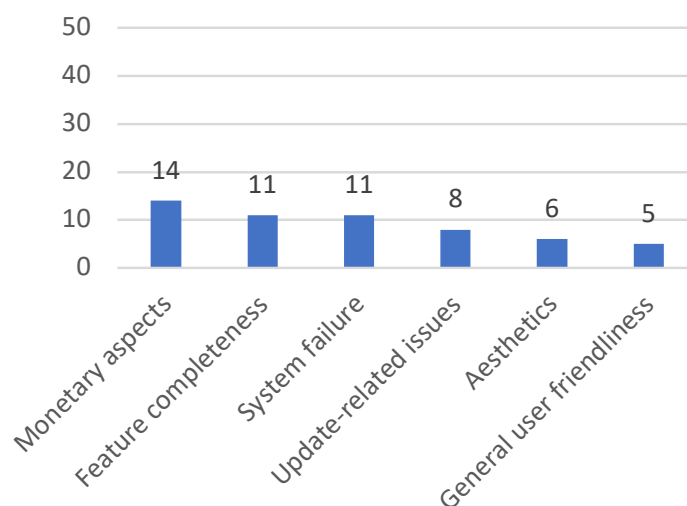


Figure 3.2 The frequency of the other categories mentioned by the users

they were not able to understand the origin of the problem and thus could not alleviate or avoid it. A Headspace user displayed their frustration with this group of problems by saying “Recently there’s an issue on the app where sleep music will stop, usually right around 5 minutes, and when I come back to the app it just says ‘loading...’ I have to close out and relaunch, and hope my attempt to relax and take a mental health break isn’t thrown off too much for the last few minutes of my break.”

(4) Update-related issues. Users raised concerns about the fact that some features or functionalities no longer worked after an update. For example, a user reviewing GoogleFit mentioned how “Every time Google Fit updates the app stops working on my Suunto 7 watch, and on my phone.

(5) Aesthetics. Users tended to like applications that had a pleasing aesthetic and a good-looking user interface. For example, a Headspace user mentioned how they like it in part because it has “such soothing colors’ . Conversely, some users had problems with bland and unexciting user interface designs. For example, a WebMD reviewer mentioned that they do not like it much because “it looks very plain”.

(6) General user-friendliness. Users liked applications that were easy to use and easy to follow. They preferred applications that allowed them to reach all of the necessary information smoothly and with speed. They also appreciated when the applications were able to run without putting much pressure on their device and hampering the function of other applications running in parallel.

3.4 Discussion

Through our analysis of user reviews of smart health apps, we found that the most frequent concerns on the smart components of these apps from the users' perspectives are accuracy, customizability, and convenience of data input. Additionally, users also mentioned other concerns in their reviews, such as monetary aspects, feature completeness, and system failure. Moreover, users are more frequently focused on the smart components (coded 181 times in our analysis) in their reviews of the apps than the other aspects (coded 55 times). Together, our results offer the following insights.

First, users tend to focus more frequently on customizability than adaptability in the reviews. In our analysis, we considered customizability as the ability of the app to give users the freedom and control of changing the behavior of the smart component, while adaptability was the ability of the app to automatically change its own behavior according to user data input. Our results indicate that users seem to pay more attention to the ability to gain control over the app, indicating their awareness, and perhaps preference, for direct manipulation over full automation. Another possible explanation is that the user experience of adaptability is often “seamless”, and thus difficult to pinpoint.

Second, there is a significant focus on the trust issue of the smart components in the reviews, reflected in four codes: accuracy, health efficacy, transparency and clarity, and scientific reliability. These codes covered about half of all the coding of the smart component concerns. Jacovi et al. [38] categorized the causes of trust of a user to an AI system as two types: (1) intrinsic trust that relies on the “observable decision process” of the AI component and (2) extrinsic trust that relies on empirical evaluation of performance. Our analysis revealed that users of smart health apps tend to focus on extrinsic trust (i.e., accuracy, health efficacy, and scientific reliability concerns), rather than intrinsic trust (i.e., transparency and clarity concerns). Interestingly, the three extrinsic trust concerns demonstrate a direct mapping to the Product Performance, Satisfaction with Previous Interactions, and Reputation aspects in Johnson and Grayson’s model [39] that contributes to the cognitive trust of customers in a service provider.

Third, in the user reviews we did not find direct discussion on many other important aspects such as safety, security, privacy, and fairness. While we do not know exactly why this is the case, there are several possible explanations. It could well be due to the reason that the users do not consider these aspects as important or lack awareness of them. Or, it is because the reviews only reflect the opinions of the users who are willing to use the apps, and who have accepted the possible risks related to these aspects. Further study is needed to investigate

these factors.

3.5 Conclusion

In this phase of the study, we conducted a thematic analysis on 120 user reviews of six mobile health apps that include a smart component to understand the most prominent concerns from the users' perspective of this type of application. Our Findings indicate that user reviews of these apps are frequently focused on their smart components; particularly, accuracy, customizability, and convenience of data input are the primary concerns. While many user reviews are focused on the trust relationship between the users and the application, very few touch on other aspects such as privacy and fairness. These findings will inspire further investigations towards guidelines that support the design of smart mobile health apps to fully achieve their potential.

Our analysis of user reviews for six mobile health apps with smart components highlights trust as a primary concern. Users prioritize accuracy, data customizability, and convenience of data input in building trust with these apps. Enhancing the trustworthiness of smart health apps is crucial to unlocking their full potential. In the next phase, we will delve deeper into trust dynamics, examining important characteristics of trustworthy smart health applications from users' perspectives. Additionally, we will explore current users' perceptions regarding the trustworthiness of these apps to inform our design guidelines and create a better understanding of the users' experiences.

CHAPTER 4 SECOND STUDY: INVESTIGATING THE ESSENTIAL CRITERIA OF TRUSTWORTHY HEALTH APPLICATIONS BY APPLYING SPECULATIVE DESIGN

4.1 Introduction

We discovered the lack of trust as a serious problem for smart health application users in the previous stage of this thesis. To address this issue we aimed to study user perceptions toward the trustworthiness of smart health applications. Trust is a complex and multilayered phenomenon and the many factors leading to its buildup have long been a subject of investigation and debate in multiple areas of study such as philosophy, anthropology, and psychology. For the purposes of this study, we used a well-established definition of trust by Simpson [22] that states “We trust based on an assessment of the other party’s trustworthiness. There is also a willingness to render ourselves vulnerable. An object or person is, in turn, considered to be trustworthy if it has [40] [41] *ability*, *benevolence*, and *integrity*. Building on this foundation and considering that interpersonal trust and trust in technology have some fundamental differences in the way that they are formed, we turned to previous work studying trust in smart technologies specifically. Toreini et al. [1] explored the trust foundations of artificial intelligence and human interaction. They established four criteria that are essential for an AI-based system in order for it to be considered trustworthy. These criteria are as follows:

- **Fairness:** This attribute applies to systems that aim to provide a fair and unbiased service for all of their users.
- **Explainability:** This attribute applies to systems that aim to explain the inner workings of their methods and functions in simple and easy-to-understand terms.
- **Audibility:** This attribute applies to systems that allow third-party and sometimes regulatory organizations to challenge and verify their quality and functionality.
- **Safety:** This attribute applies to systems that provide safe and low-risk products to the users that guard and protect their data and privacy.

We leveraged this framework for our study regarding smart health applications and aimed to investigate whether they are effective indicators when it comes to user trust.

Smart health apps are evolving at a fast rate in a multitude of manners. So, it was particularly important to investigate the prospect of trustworthy applications in the future and the factors that could affect their levels of trustworthiness. Additionally, many of the previously discovered hurdles to users’ trust in smart health applications stemmed from the inaccuracies and miscalculations present in the apps, relating more to technical limitations than to user experience design practices. The solution to these issues is outside of the domain of user experience design, moreover, many of these technological capabilities are developing and evolving quickly making their corresponding limitations to be also fluid and fast-evolving. This quality makes it difficult to build design guidelines that transcend the current technological inaccuracies and limitations. In this thesis, however, we aim to explore the broader design practices that could lead to the development of trustworthy smart health applications regardless of the current technological shortcomings and limitations. To nullify the effect of technological limitations and account for the rapidly changing landscape of smart health apps, we use a *speculative design* approach. This approach creates a controlled environment, allowing us to focus on the core elements that foster user trust, without being tied to technological limitations.

Different from traditional design approaches such as problem-solving design or functional design, speculative design explores the social, cultural, and ethical implications of emerging technologies by envisioning futuristic or alternative scenarios [11] [42]. It provokes critical thinking and reflection on the broader implications of technological and societal changes.

Given its unique characteristics, speculative design is particularly useful for investigating potential future paths for technologies and design practices. By looking at futuristic or alternative societies where current limitations do not apply, it allows us to zero in on the user experience design aspects of smart health applications.

In this study, we aim to assess the applicability of Toreini et al.’s four criteria of trustworthiness [1] in smart health applications using speculative design. Through this approach, we explore potential future design directions to enhance trustworthiness and benefit users, while considering their perception of the transition. We also recognize the importance of understanding societal trajectories impacting technology development and users’ trust implications. Our overall objective is to shed light on characteristics leading to user-perceived trustworthiness in smart health applications and understand how it reflects in the societal environment. Ultimately, we seek valuable insights to design user-centered experiences that foster trust and address evolving needs and expectations.

To work toward this goal, we first developed a series of society descriptions detailing the speculative environment of futuristic societies and their approach to the design of smart health

technologies. For each society, we then designed corresponding smart health application prototypes that we think the society will value. The society descriptions and the app prototypes were iterated based on preliminary user feedback. Then we conducted a user study to gather users’ opinions and their reflections on the society descriptions and the app prototypes. During the study, we explored users’ perceptions of the potential of each criterion for creating trustworthy smart health systems. We analyzed and compared the results to probe into the importance of each of the four criteria and their relationship with each other.

4.2 Iterative Speculative Design of the Prototypes

4.2.1 Society Design

As discussed, we chose the exploratory method of speculative design to navigate the rapidly changing technological capabilities and limitations of smart systems and to investigate the broader development of trustworthy smart health applications. The trajectory of smart health applications, as well as smart systems in general, is undeniably influenced by the overall development and societal trends [43] [44]. Thus, envisioning the future equivalent of a smart health application necessitates consideration of the broader societal context in which it operates. To initiate the design process, we developed several fictional societies, each evolving in a way that allowed for varying degrees of *fairness*, *explainability*, *audibility*, and *safety*. Each society description was carefully crafted to provide sufficient background information about the fictional state of the society and detailed insights into its health systems and smart health systems while still encouraging imagination and creative interpretation, crucial aspects of speculative design.

Toreini et al.’s framework [1] only provided very high-level definitions of the four trustworthy criteria. In order to apply these definitions to our speculative design process, we refined these definitions and considered the meaning of both extremes of each criterion in our design context:

- **Fairness:** technologies focused on the detection or prevention of discrimination and bias in different demographics and abstained from generalizing toward a dominant group.
Low fairness: Not addressing individual differences. Unfair to an under-resourced population
High fairness: Sensitive to discrimination and bias. Address individual differences.
- **Explainability:** technologies focused on explaining and interpreting the outcome to

the stakeholders (including end-users) in a humane manner.

Low explainability: No one knows how the technologies work; people just blindly follow the recommendations.

High explainability: People understand the suggestions/recommendations that the technologies provide and follow them with discretion.

- **Audibility:** technologies focused on enabling third parties and regulators to supervise, challenge or monitor the operation of the model(s).

Low audibility: No one knows what the companies are doing.

High audibility: The company needs to give sensitive information when requested.

- **Safety:** technologies focused on ensuring that the system is medically safe (data is correct and the app is safe to use, not harmful to users) and the operation of the model works as intended in the presence of active or passive malicious attackers.

Low safety: The technology is not necessarily medically safe or accurate. Attackers can easily gain access to technologies, and thus personal data, etc.

High safety: There is a reasonable level of medical safety and using them carries little risk to users both in terms of potential medical harms and in terms of possible data breaches and privacy transgression.

We started by creating 16 fictional societies that encompassed all possible combinations of the extremes of the four criteria (e.g., low-low-low-low, low-low-low-high, etc.). To emphasize the significance of each criterion and maximize differences between the societies, we intentionally focused on the extreme ends of the spectrum. This meant that each hypothetical society exhibited a particular criterion in either very high or very low degrees. After inspecting those fictional societies, we found that the 16 society descriptions shared numerous similarities, with some scenarios appearing highly improbable and far-fetched. For instance, envisioning a world where smart systems were incredibly fair, transparent, and easily understood, yet remarkably unsafe, posed significant challenges. To address these issues and streamline the user study, we undertook a trimming-down process.

During the trimming-down process, it became evident that *Fairness* and *Explainability* were closely related – a fair society necessitates providing information to reassure individuals about the system’s fairness and suitability; and similarly, a transparent society would naturally produce diverse and unbiased products, as its citizens can access relevant information and make informed decisions. Likewise, it was apparent that a robust system of checks and balances by regulatory bodies and trusted organizations (i.e., *Audibility*) leads to safe technologies (i.e., *Safety*). Developing and maintaining genuinely safe products without such checks is highly

challenging. These pairings guided the selection process, resulting in the development of four final fictional societies.

The four selected societies were specifically designed to explore the effects of the four trustworthiness criteria and compare the importance of each pairing. Here are the descriptions of the final four selected societies, labeled with the considerations of the four trustworthy criteria:

Society A - Low Fairness - Low Explainability - Low Audibility - Low Safety

In society A, most people are treated like homogeneous bricks without much individual worth. The continuous growth of the world population, the decline of natural resources, and the effects of global warming have rendered most people expendable cogs in the machine. The governments in most countries prioritize low cost and minimize the burden on the health sector through solving most common problems with general solutions. The health services are working based on the data received from the most populated groups and cultures. As a result, members of the more underprivileged groups are mostly ignored. Most people receive general advice through public websites and applications accessible through mass-produced health check devices with medium to low sensitivity sensors worn on their wrists waist or around their neck. The general effectiveness of these methods is never checked. And as long as there are no catastrophic results these technologies continue to be used. People can access their files and info on any personal computing device but they can not check the reasoning behind a suggestion or health tip. As a result of the emphasis on keeping the cost down, the vulnerability toward nefarious online attacks is high and it is not rare for the public data to be leaked. Additionally, the government can access the health data of citizens without their consent. The mass solution to health problems, however, keeps the overall process financially cheap.

Society B - Low Fairness - Low Explainability - High Audibility - High Safety

After the globalist centralization of the world government and decades of uniform education provided to most people, many of the individual and cultural differences have started to fade. Groups that do not conform with what the society believes are marginalized and their interests are discarded when it comes to making public policy decisions. To reflect on this worldview, each citizen receives tracking devices that record their exact physiological data since birth. They are required to wear their wearable devices at all times. All of the services and devices used to monitor and improve public health have to pass a rigorous examination process by government institutions so that their functionality is scientifically proven. On the

other hand, the diversity of the available products is limited – new and exciting products can not enter the market before going through the lengthy process for approval. Additionally, people receive tips and recommendations, and notifications constantly. However, they can not know the exact reason behind these suggestions or check the track record of the methods offered to them. This forced ignorance can cause people to miss better options for health services and solutions. The public hospitals and medical centers have a mixture of diagnostic supercomputers which are directly connected to the database storing the health information of people in their district and some human healthcare workers, who are also important for solving more severe health issues. As a result, each individual’s recommended health program is made to minimize any danger or harm. Moreover, the gathered health data is highly guarded and there are multiple protections in place to prevent a data attack. However, as a product of the overall philosophical approach of society, an individual’s physical and mental characteristics are not part of the health information being recorded. Their personal preferences are never taken into consideration. As a result, each individual is confined to the highly detailed but general list of health services provided to them.

Society C - High Fairness - High Explainability - Low Audibility - Low Safety

In Society C, people are used to personalized AIs that feel like a friend. It’s very important for them that they have their “friends” built in the specific way that they prefer. With the emphasis on multiculturalism and representation of each group, most service providers try to have a wide variety of services for different groups and demographics, genders, ethnic groups, as well as sexual orientations and identities. Companies also prioritize having a diverse development and design team. This emphasis however sometimes can turn into catering to a marketing strategy – some good products can be rebuffed just because they do not look diverse enough. Basically, each user in Society C is the master of their own health journey. Thus, health systems emphasize providing users with abundant customization options. On the downside, giving so much control and responsibility to the users causes some citizens of Society C to fall into the trap of believing themselves to be professionals capable of making every health decision on their own. In practice, this can prove to be quite dangerous. Similarly, the lack of oversight means that sensitive health data can move through very unsafe channels which are extremely vulnerable to attacks. Moreover, as so many health-related decisions are made by the users and the process is distributed through a very large and open network, any form of regulated testing and universally reputable certificates are impossible to employ properly.

Society D - High Fairness - High Explainability - High Audibility - High Safety

In Society D, automation has grown to take over a substantial part of human responsibilities. Here, different and competing automated services are offered to people in any domain, although some human oversight in key and sensitive domains such as health care is still deemed necessary. This society offers each individual a personalized and detailed set of services curated just for their habits, personal traits, and preferences, emphasizing the diversity of users in the design process of apps and services. It is important for each health service provider to prove that they used a diverse and reflective set of users in the testing process and employed the help of relevant experts to make sure that the service provides valid information to each demographic. Although beneficial to avoid injustice and bias, this approach can make the testing and development process harder and can push functional products that are not specially designed to emphasize diversity out of the market. The users are free to check the reasoning behind a decision or suggestion; they can even look at the inner workings of an algorithm. They also can see the history and past data for the available health methods which can help them make a more informed decision. The access to personal data, however, is limited to what has been explicitly granted to them by the users. The vigorous safety testing ensures that the technologies are capable of protecting users' data. Only AI systems passing the safety tastes and medical trials are allowed to offer health services. Users still can apply some personal preference to the systems and choose between the variety of available options. But as the general census is inclined to opt for safer, more well-known options, and due to the expensive trials and medical technologies required to perform specialized practices, most health services are offered by a number of big companies.

4.2.2 Initial Prototype Design And Design Iterations

Upon creating the descriptions of the different societies, we proceeded to develop four high-fidelity prototypes that closely resembled smart health applications in terms of visual and aesthetic properties. This approach aimed to provide users with a familiar interface, allowing them to concentrate on the four criteria represented in the prototypes and their overall user experience. By utilizing a design that is familiar to the user, we create a possible-plausible effect that is essential to speculative design [42] [11], allowing users to reflect on alternative and futuristic values based on their understanding of the current reality. However, it's worth noting that the prototype corresponding to Society C deviated slightly from this approach to represent the individuality and occasional quirkiness of the products within that society; this decision is made because being unique and sometimes unconventional is an integral characteristic of the systems in Society C. This diversion however, was still minor and mostly

reflected on the aesthetics and did not affect the way that the users interacted with this prototype. All of the prototypes were developed using *Canva*¹.

To avoid introducing bias, the prototypes did not feature any explicit visual indicators of their corresponding societies or the levels of fairness, explainability, audibility, and safety they intended to represent. Instead, these qualities were subtly conveyed through the functionalities of the prototypes, the information provided, and the user experience they offered. *Fair* systems presented personalized and flexible options tailored to users' needs and characteristics while allowing for choices. *Explainable* systems made the reasoning behind their suggestions and the accuracy of their methods clear and understandable. *Audible* systems highlighted the tests and checks conducted to ensure product effectiveness. Lastly, *safe* products refrained from recommending risky or unhealthy suggestions or services while safeguarding user privacy.

Each prototype depicted a different health-related function, although participants were informed that most prototypes were capable of performing all regular health-related functions, such as diet, fitness, activity tracking, mental health, and diagnostic tools. This approach served two purposes. Firstly, it diversified the range of activities and scenarios encountered by the users, expanding their imagination and reflection regarding the applications, in line with the goals of speculative design. Secondly, it prevented participants from fixating on the minute technical details of how each prototype performed a similar function and encouraged them to focus on the broader and inherent differences in the design ideology of each prototype.

After the first set of prototypes was created, we conducted two rounds of preliminary user studies to iteratively refine the prototype designs. Three participants aged from 23 to 25 participated in those user studies, including one male and two females. During the study sessions, the participants were first asked to explain their background experiences with smart health applications and answer a number of demographic questions. The next step of the study asked participants to work with the four designed prototypes, finish a series of tasks and share their opinion regarding the perceived trustworthiness of these prototypes. Lastly, participants were presented with the four fictional society descriptions. They were then asked to read these descriptions and share their thoughts about the societies. Specifically, they were to consider how suitable these societies appear for the development of trustworthy smart applications. The sessions were recorded and their responses were then analyzed to extract the major problems of the study and refine the prototypes. Based on the participants' feedback, we made various changes to the prototypes. We describe two major changes below.

¹<https://www.canva.com/>

The first major change is related to the color schemes used. Initially, the prototypes featured distinct color themes intended to emphasize the individuality of each prototype. For instance, the first prototype had a grey and black color theme to reflect the austere nature of the society it represented. However, during the first round of pilot studies, the first participant pointed out that these color choices could introduce bias and unduly influence their perceptions of the prototypes. Taking this feedback into consideration, we made the decision to transform all the designs to incorporate green, blue, and white colors. These colors are commonly associated with health and healthcare, ensuring a more neutral and unbiased visual presentation across all the prototypes. As an example, Figure 4.1 shows the old color theme of Prototype One which was deemed too dull by the participants.

The second major change was related to the design limitations that inadvertently appeared as intentional design choices. For instance, prototype 3, representing Society C, included a recipe recommendation system, where users were able to get recipes based on either a health-oriented plan or a plan based on the foods they liked. This prototype was originally perceived by two participants as restrictive due to its display of only one recipe for each plan. This experience originated from the prototype’s incomplete and non-functional nature, rather than a deliberate design decision. To address this issue and provide a more comprehensive user experience, we added more pages and functions to the prototypes to offer a broader range of options and enhance overall usability. These enhancements aimed to ensure that users could experience a more realistic representation of the potential application’s capabilities. Figure 4.2 shows the page added to Prototype 3 which presented the participants with more recipe options.

Additionally, we also found that the participants had a difficult time examining the prototypes if they did not know the context of the possible societies. This observation inspired us to reverse the order of presentation of the prototypes and society descriptions in the main user study that we will describe in Section 4.4.

4.3 Final Prototype Design

Below is a short description of the final versions of each of the prototypes, an overview of all of the pages of the prototypes is available in the appendices.

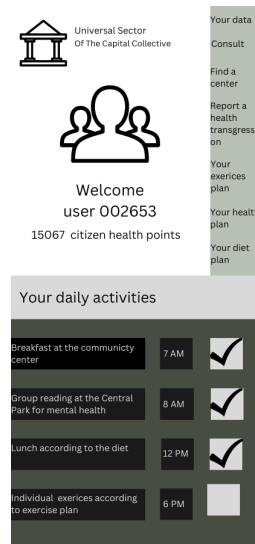


Figure 4.1 The old color theme of the prototype 1

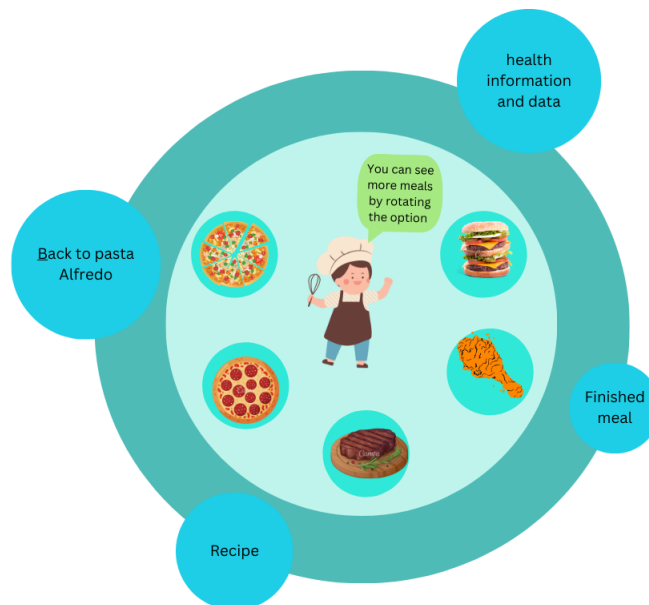


Figure 4.2 Adding additional recipe options to prototype 3

Prototype 1 corresponding to society A with low fairness, low explainability, low audibility, and low safety

Prototype 1 served as an activity recommendation and tracking system; Figure 4.3 demonstrates its first page and some of its functionalities and limitations. On the main page, users could view their citizen health points, representing rewards earned upon successful completion of recommended activities. Users lost points if they were unable or unwilling to finish an activity. It was explained to the participants that these points contributed to their social standing within Society A. The main page provided users with daily activity instructions.

The daily recommended activity section allowed users to mark completed activities with a ticked box, while missed activities were identified with a cross mark. Activities that remained unfinished had an accompanying empty tick box. Clicking on an activity led users to an activity page with instructions on how to perform the task. Although the prototype focused on the daily exercise section for the user study, participants were informed that it aimed to represent a comprehensive daily health-related activity recommender and tracker. Other sections were not developed due to time constraints.

For individual exercise sessions, a video demonstrating the recommended exercise was provided to users. The audio was intentionally removed to add difficulty. Two options, “Successfully finished the exercise” and “Could not finish the exercise,” were presented. Selecting the first option resulted in users being informed that they did not properly complete the exercise according to the tracking data, without offering a solution or further explanation. This led to no points gained or lost. Choosing the second option notified users of lost health points, again without providing a solution or justification.

The lack of *fairness* was portrayed through the impersonal and cold interface treating citizens as mere numbers, as well as the inflexible nature of recommended activities that disregarded personal preferences. To underscore this unsuitable nature, all users were recommended a relatively difficult daily exercise with inadequate instructions. The absence of *explainability* was demonstrated by not providing information about the basis for suggestions or recommendations and informing users of exercise failure even when they selected the successful completion option, without any further explanation.

The lack of *audibility* was evident in the absence of information, reassurance, or certification regarding any tests or checks conducted to ensure the functionality of the product. Similarly, the lack of *safety* was depicted through the absence of privacy information and ambiguous tracking without user consent. Additionally, the difficult recommended exercise, regardless of users’ medical and physical conditions, highlighted the disregard for safety.

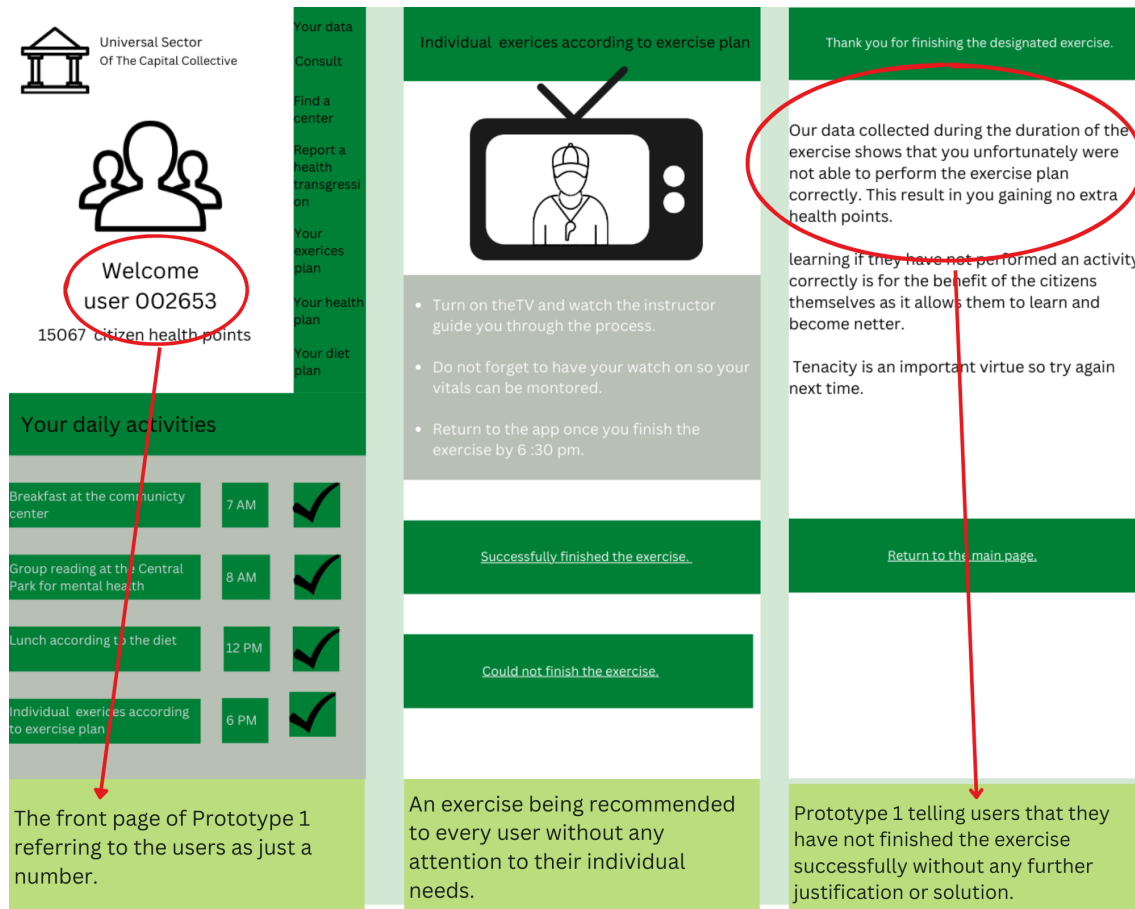


Figure 4.3 Prototype 1

Prototype 2 corresponding to the society B with low fairness, low explainability, high audibility, and high safety

Prototype 2, representing Society B, was a comprehensive health tool; Figure 4.4 showcases a snapshot of its initial page and a selection of its functionalities and issues. The first page featured sections for daily activities, checkups, diet, and general health tracking. While the daily activity section was explored in the prototype and user study, participants were informed that the hypothetical application it represented included all sections.

Users had the flexibility to access any recommended activity by clicking on it in the daily exercise section. They could then view instructions for the activity and begin it. For example, in the daily walk activity, the prototype informed users about the optimal time for their walk and allowed them to schedule it within the specified time slot, with a suggested distance and path. Upon completing an activity, such as a daily walk, users were congratulated, and the activity was marked as finished.

The high level of *audibility* in Society B was demonstrated by informing users that the application had been tested, approved, and endorsed by medical professionals. The emphasis on *safety* was evident through assurances regarding the privacy of users' data and the mention of vital signs being monitored to prevent any potential harm. Furthermore, the app informed users that their walk had been scheduled during the most medically beneficial time, adding to the medical safety aspect.

However, the lack of *fairness* was depicted by the provision of a strict and inflexible schedule that did not consider individual needs. Despite claiming to be medically optimal, it provided no flexibility or choice for users with time or other limitations. The absence of *explainability* was conveyed by not providing users with explanations as to why certain activities were suggested or why they were unable to choose an alternative time for their daily walk. While it was mentioned that their vital signs were being recorded, no further information or conclusion was provided, highlighting the lack of explainability once again.

Prototype 3 corresponding to society C with high fairness, high explainability, low audibility, and low safety

Prototype 3, representing a fictional application in Society C, featured the most visual differences compared to the other designs; Figure 4.5 highlights the first page of this prototype, showcasing a range of its functionalities as well as its inherent limitations. This intentional divergence of visual design aimed to emphasize the individual and creative nature of the products in the society it represented, without guaranteeing their proven functionality. The

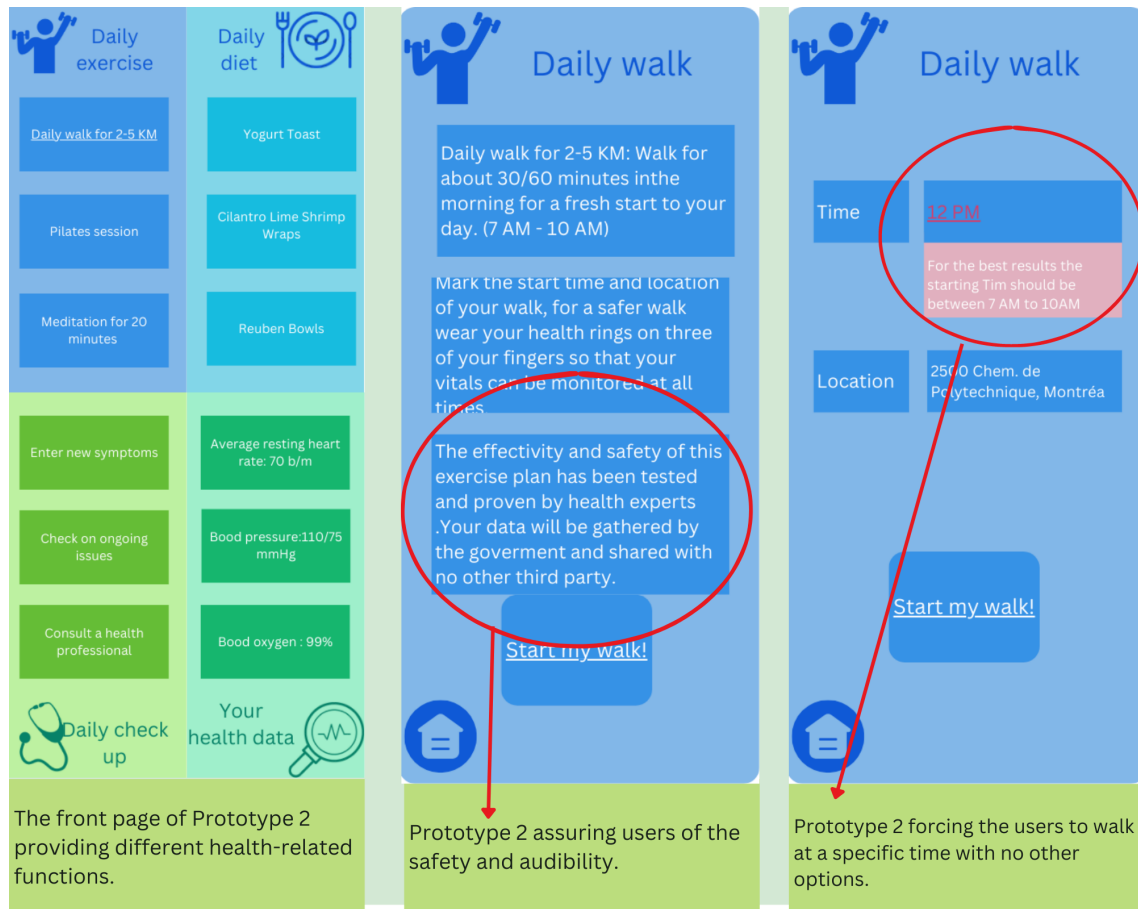


Figure 4.4 Prototype 2

prototype served as a complete health-related activities tool, with visual elements resembling a circular device. The smaller circles projected outside the main device, representing the user's daily recommended activities.

The user's current activity (e.g., a recommended lunch recipe) is highlighted by a dish. However, users had the option to select another activity by circling through the icons. The diet and recipe recommendation section was depicted in the prototype and user study, while participants were informed that it was just one part of the application's functionalities.

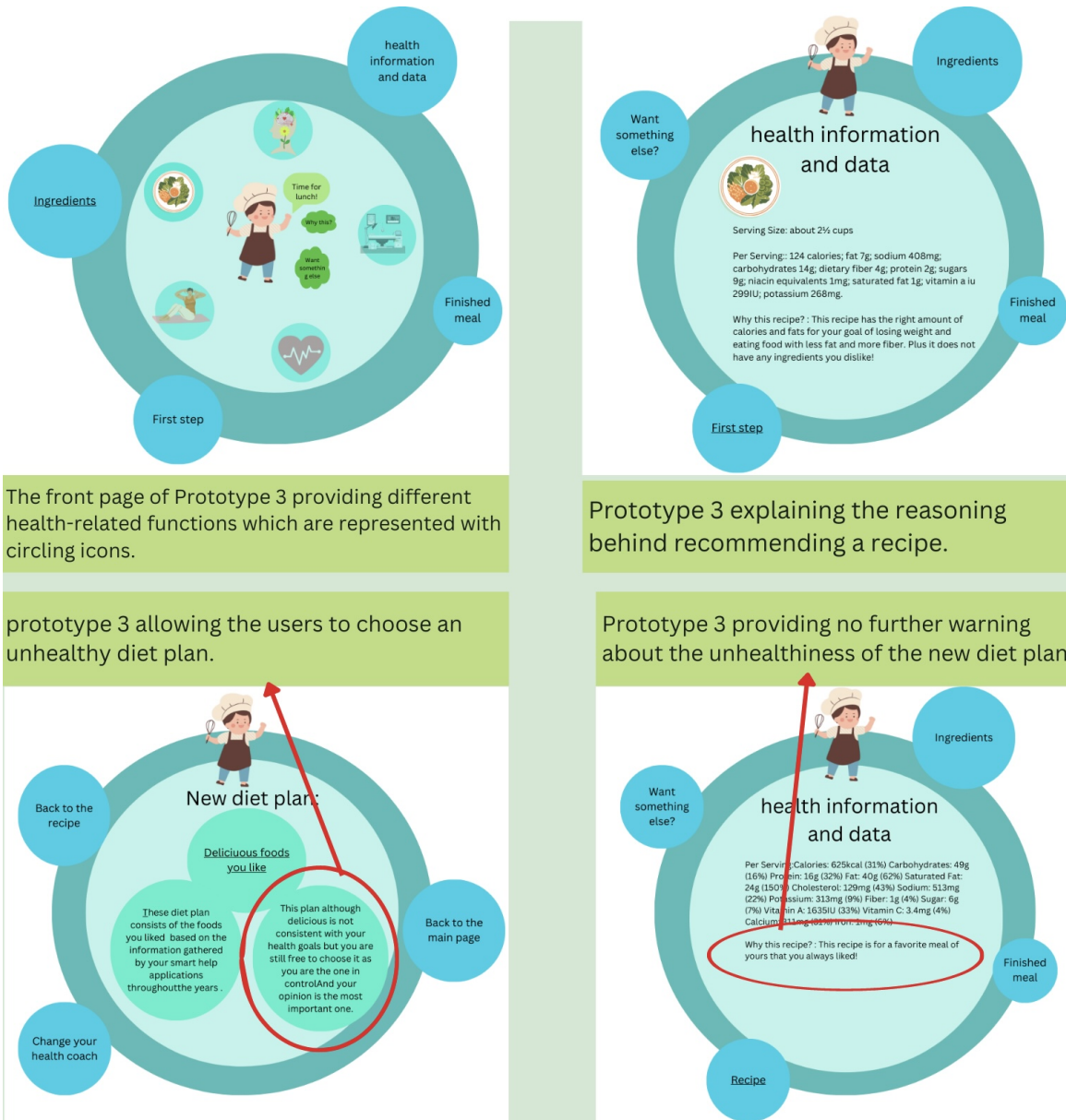
Recipes were provided to users, including health information and the reasoning behind their recommendation. Users could select another recipe with similar nutritional value or completely change their diet plan to different potential plans. This included a plan that disregarded nutritional facts and recommended each user's favorite food, although users were warned that this plan lacked medical input and did not align with their health goals. No further warnings were given to users, and they were allowed to choose and follow the plan. The prototype offered step-by-step recipes, ingredient lists that only included disliked ingredients, and a picture of the final meal.

High *fairness* was expressed through personalized suggestions, including diet plans and recipe recommendations tailored to individual users' data and goals. The prototype also considered users' dietary and monetary limitations when suggesting recipes, ensuring they did not exceed their budget or include ingredients they could not consume. Users were provided with choices, allowing them to modify plans according to their preferences. High *explainability* was depicted by providing explanations about why specific suggestions or recipes were offered, as well as the benefits and risks associated with each plan.

However, the lack of *audibility* was apparent due to the absence of assurance or certificates regarding tests, permissions, or checks conducted to ensure the system's functionality. Similarly, the low level of *safety* was conveyed through the absence of visible privacy measures and the system's suggestion of an unhealthy option, such as an unhealthy diet plan, when users opted for it, potentially compromising the system's medical safety.

Prototype 4 corresponding to the society D with high fairness, high explainability, high audibility, and, high safety

Prototype 4, corresponding to Society D, served as a comprehensive tool encompassing various health-related functions such as diet, exercise and tracking, mental health, and diagnostic tools. The prototype further explored diagnostic tools, which were utilized during the user study. Similar to the previous prototypes, participants were informed about the existence



of other health-related functionalities within the application represented by this prototype. Figure 4.6 showcases its initial interfaces.

In this prototype, the scenario involved obtaining a diagnosis for a minor health issue, such as a relatively mild headache. Users had the ability to review potentially related health information gathered by the prototype. They could investigate how this data was collected and make changes if they found any inaccuracies. Users were then prompted to enter additional symptoms and provide relevant information. Subsequently, the prototype provided them with a calculated diagnosis, along with information on how the diagnosis was reached, recommended treatments and their effectiveness, other possible conditions that could be responsible for their symptoms, and guidance on when to seek further medical attention.

The high level of *fairness* was demonstrated by offering personalized recommendations, services, and diagnoses based on users' individual needs and conditions. Users had the flexibility to choose different options or modify settings within medically safe parameters. Additionally, the diagnosis process relied on the users' individual data and input. Greeting users by their name and displaying their picture added a sense of familiarity and personalization. The high level of *explainability* was conveyed by providing reasoning and explanations about the data collection process, result generation, and the probability of accuracy. For the depicted diagnostic section, information was provided about the effectiveness of the recommended treatment.

The prototype ensured a high level of *audibility* by reassuring users that the system had undergone testing and received approval. Similarly, the high degree of *safety* was highlighted by guaranteeing user privacy protection and storing only the information that users had previously authorized. The involvement of medical professionals and the system's high accuracy level contributed to the medical safety aspect. The prototype also included warnings for situations where users needed to seek medical attention from real medical centers, further enhancing safety levels.

4.4 User Study Methods

After finalizing the design of the society descriptions and prototypes, we conducted a user study to assess the perceived trustworthiness of each society and design. The user study is approved by the ethics review board of Polytechnique Montreal.

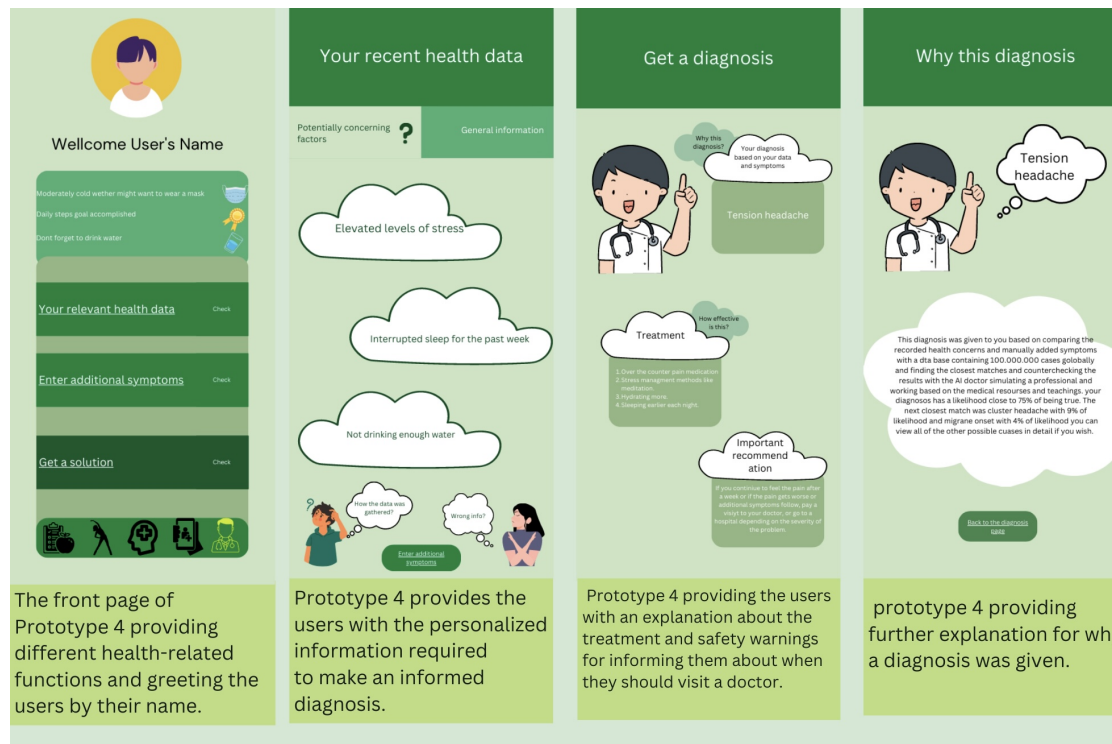


Figure 4.6 Prototype 4

Participants

Six participants were recruited for the user study. All participants had a high level of technology literacy and familiarity with smart systems, allowing the focus of the study to be on the design differences influencing trustworthiness rather than explaining the fundamental elements and functions of smart health systems. Participants were recruited using a snowballing method. Table 4.1 demonstrates their demographic and background information.

Study Process

Each study session lasted approximately 90 minutes and was conducted online over Zoom. All study sessions were recorded, with the recordings stored securely and privately. Participants were informed in advance about the recording and provided consent. The study session comprised three main parts.

In the first part, participants were asked about their prior experiences with smart health apps, focusing on the elements they found smart in those apps, positive experiences where they trusted the apps and negative experiences where they lacked trust. Participants were also asked about broader issues and design practices influencing their positive or negative

Participant	Gender	Age	Education	Location
P1	Female	36	Ph.D Student	Montreal
P2	Male	23	Master's Student	Montreal
P3	Female	28	Master's Student	Montreal
P4	Male	28	Ph.D Student	Montreal
P5	Male	24	Master's Student	Montreal
P6	Non-Binary	25	Master's	Montreal

Table 4.1 Profile of the participants of our user study

experiences. Additionally, participants shared their general thoughts on the trustworthiness of smart technologies in our world today and their perceptions of what is needed in an ideal society to create trustworthy smart applications. The purpose of these questions was to evaluate participants' previous experiences and expectations and to compare them with the prototypes. Additionally, the questions aimed to analyze participants' positive and negative experiences in relation to the four criteria of trustworthiness. Furthermore, the questions sought to explore the relationship between the trustworthiness of smart technologies and the characteristics of the world in which they are developed and utilized.

The second part of the study involved participants reading a document containing the descriptions of the four fictional societies one by one. Participants answered questions related to each society, without being provided information about the four criteria associated with each society. The questions inquired about participants' overall opinions of each society, their perception of the society's characteristics in relation to the development of trustworthy smart health applications, and similarities and differences between the society and the real world. After reading the descriptions of all societies and answering the corresponding questions, participants ranked the societies based on their perceived trustworthiness.

In the third part, participants were provided with each prototype and asked to engage with them and complete various tasks. The prototypes were given to participants in a randomized order, without informing them which prototypes represented which society. After completing each task related to a prototype, participants were asked to share their feedback on what they liked and disliked about the prototype. They were also asked to evaluate the trustworthiness of the prototype and identify the biggest similarities and differences between it and the smart health applications they had previously used. Furthermore, participants were asked to map the prototypes to the societies they believed they were related to. Participants had the flexibility to associate multiple societies with a single prototype or multiple prototypes with a society. This question aimed to examine whether the prototypes were able to accurately represent the societies to a reasonable degree.

Data Analysis

We conducted a thematic analysis on the qualitative data gathered from the study sessions. All study sessions were transcribed. The transcriptions underwent meticulous scrutiny, with data segments being systematically labeled and organized into overarching themes. These themes were carefully aligned with the research objective, specifically focusing on participants’ reflections on the trustworthiness of smart health applications and the influence of the four criteria on their perceptions. Additionally, we established other thematic groups to delve into participants’ perspectives on trust and trustworthiness, considering their prior experiences with smart health apps, the descriptions of the societies, the prototypes, and any other pertinent factors that could have impacted their assessments of trustworthiness. The identified themes were subsequently utilized to extract the results.

4.5 User Study Results

4.5.1 Previous Experience With Smart Health Apps

The results of the first part of the study showed that, as expected, the users had a wide range of experiences with different groups of smart health applications. The most popular smart health apps were the activity trackers (used by all six participants), followed by workout planners (five participants) and calorie calculators (four participants). Among the activity trackers, pre-installed default fitness applications that come with the phones (e.g., Apple Fitness and Samsung Fitness) were the most popular and used by four participants. Other workout apps like *Strava*, *keep*, *Adidas*, and *Nike*, as well as other calorie counter apps like *Karafs* were also used by the participants. One other participant mentioned using a period tracker app called *Flo*. None of the participants mentioned using a smart diagnostic app. All participants recognized the presence of smart components in the apps, specifically referring to either the recommendation system’s capability to provide informed suggestions or the apps’ ability to utilize data, sensors, and pictures, and accurately calculate some health-related data. For instance, an example of the latter type of intelligence was observed in the apps’ capacity to scan food and calculate its calorie content accurately.

Participants shared generally positive experiences with the applications. They attributed these positive experiences to the applications’ ability in showing them accurate information and providing them with a personalized plan or suggestion. For example, P2 emphasized the importance of this flexible and personalized quality by complementing an app by stating: “*I mostly like to run on flat areas, so it recommended two or three routes that were all flat.*” P5 also shared: “*So the, the good thing that I, I like about this application is the recommendation*

system, like based on my activities.” They, however, did have some negative experiences and instances where they could not trust the applications. These instances were usually the result of inaccurate information provided by the apps, such as showing an obviously wrong heart rate, being unclear about their methods, or privacy projections and data gathering methods.

The participants highlighted several additional important factors when assessing the trustworthiness of the applications. Accuracy, clear explanations regarding app functionality and data gathering, and the safety of user data and privacy were emphasized. P highlighted the importance of accuracy when sharing their experience with a tracking app by stating “*I can trust it because I know the accuracy is pretty good and it does not give me wrong information*”, on the other hand, P3 shared their distrust about the IOS tracker resulting from its vague data gathering methods by saying “*they collected a lot of information about my sleep duration at night, for example. And, I didn’t know that it’s collecting this information but when I go to the health app, I can see a history of my sleep duration for the past month! And it’s really not a good experience when you realize that someone has been collecting your information and you didn’t know about it.*” Furthermore, the participants recognized the significance of “*regulations*” (P5) and the involvement of health experts. They also expressed a desire for less emphasis on “*marketing claims*” (P1) and monetary motivations, among other noteworthy factors that influenced their assessment of the applications’ trustworthiness.

When asked about a perfect world where smart health applications will be trustworthy, the majority of the users mentioned the importance of having some type of regulation or audit process in place to make sure the applications are safe and/or effective. They also mentioned the need to use “*health experts*” (P3) during the development of the application and make the users feel like they do matter instead of focusing on just “*money, money, money*” (P1). Privacy and the vagueness of what data is collected were also mentioned by the majority of the users. As an example, P2 said “*I think only making everything clear would be the first step to just let them (users) know what type of data is being monitored*” when asked about the hypothetical trustworthy society.

4.5.2 Evaluation Of The Society Descriptions

All participants unanimously agreed that Society A is not conducive to the development of trustworthy smart health applications, while Society D was considered well-suited for such development. However, opinions regarding Society B and Society C were more diverse. Four participants found Society B to be untrustworthy, while two participants deemed it somewhat trustworthy. Similarly, two participants found Society C entirely untrustworthy, two found it trustworthy to some extent, and another two participants regarded it as trustworthy. We

report participants' perceptions of each society below.

User Perceptions on Society A

Participants generally shared a very negative view of Society A, describing it as rigid and devoid of any freedom and choice. They also expressed concern about the lack of privacy and safety measures. P3 expressed their opinion by stating: *"I really don't like to be in that society. Because of the privacy, that was one of my concerns. And the fact that the people are treated like they are, all similar."* Participants also noted how some groups that differ from the general norms, *"are ignored"* (P6), and how much the focus is on making the products cheap and universal. A quality that can mean that people's *"health is in danger"* as the focus is on making a product cheap and not on making it safe or functional, according to a P1. The participants also highlighted how the lack of transparency can worsen people's situation in this society.

The main factor appearing as a difference to the participants when comparing Society A to our society was the lack of choice and diversity. Although P5 thought that even in our own society the needs and preferences of a dominant group usually get expanded to anyone, the situation in this hypothetical society seemed a lot more extreme to all of the participants. P4 mentioned the technology being cheap in Society A as one of the few positive points.

User Perceptions on Society B

Although participants in general viewed Society B more favorably than Society A, the general overview remained mostly negative. P3 mentioned that Society B is better at protecting people's data compared to Society A and P2 noted that there seems to be some level of care for individuals through regulations and checks. However, participants still expressed dissatisfaction with this society due to its lack of freedom, flexibility, transparency, and respect for diversity. P5 described Society B as *"rigid and robotic,"* where people have limited freedom, almost resembling a dictatorship.

Transparency was a significant concern raised by multiple participants, as they felt that people were unaware of the reasoning behind the health plans provided to them. Consequently, four participants deemed Society B unsuitable for the development of trustworthy health applications. The two participants who found it somewhat trustworthy emphasized the importance of accuracy in the applications, as it could compensate for the lack of explanation. It is also noteworthy that the added regulations for the audibility and additional safety precautions lead to view Society B as even more restrictive and authoritative than Society A

and rank it lower in terms of trustworthiness.

The lack of freedom, choice, and consideration for individual needs and conditions emerged as the primary points of distinction between Society B and our own society. Participants also highlighted that our society generally does not have the same level of checks and audibility. Some participants drew parallels between situations in which specific health apps or society as a whole provide advice or commands without sufficient explanation, linking it to similarities between Society B and our society.

In summary, participants acknowledged some improvements in Society B compared to Society A, particularly in data protection and regulatory measures. However, concerns about limited freedom, lack of transparency, and disregard for individual differences remained prevalent, influencing their perceptions of Society B's suitability for the development of trustworthy health applications.

User Perceptions on Society C

Participants held mixed views regarding Society C. Some participants perceived it as a multicultural place that respects choices, while others considered it unsafe and dangerous. Negative aspects mentioned by participants included the lack of regulations and safety measures, which could potentially lead to health issues and excessive reliance on self-trust with negative consequences. P4 specifically highlighted data protection and the usability of AI as significant concerns in Society C. As a result of these conditions, two participants considered this society untrustworthy.

Positive characteristics mentioned by participants primarily focused on providing choices to individuals and accepting them for who they are, considering their unique needs and situations. P2 expressed, "*This one is better because people are choosing the AI that they're working with. So it should be something to satisfy the people.*" Participants also acknowledged this society's tendency to provide information and explanations.

The most significant similarity between Society C and our own society was the presence of multiple choices and the emphasis on diversity. Participants noted that our society has some checking systems and regulations in place, unlike in Society C. Additionally, participants highlighted that smart systems in our world are not as sophisticated and personalized as they are in this fictional society. P6 highlighted this point by stating: "*No AI as advanced and personalized exists in our world now.*"

In summary, participants had differing opinions about Society C, with some recognizing its positive aspects of providing choices and respecting diversity, while others expressed concerns

about safety and the lack of regulations. The presence of educational knowledge about smart systems was considered a factor influencing trustworthiness in this society. Participants also noted the differences in regulatory systems and the level of sophistication in smart systems between Society C and our own society.

User Perceptions on Society D

All participants unanimously agreed that Society D is the optimal environment for the development of trustworthy smart health applications, with all participants ranking it as the most trustworthy society among the options. P2 described Society D as the utopia towards which we should strive. Participants highlighted Society D's ability to offer choices, customization, and respect for individual differences, creating health services tailored to individual needs. In Society D, smart health products are developed based on people's habits, traits, preferences, and diversity, resulting in personalized healthcare for each person. Transparency and safety measures in Society D were also praised by participants. The visibility of algorithms and programs, as well as robust data protection, were cited as important characteristics that contribute to the trustworthiness of applications in this hypothetical society. Participants also noted the existence of regulations and audit processes as positive aspects of Society D.

Several participants mentioned that while our society aims to achieve the situation depicted in Society D, we have not yet reached that level. P1 expressed the view that our society is about 70-80 percent there but still has room for improvement. P6 also noted that our world lacks the same level of transparency and explainability as Society D. Participants also observed that artificial intelligence appears to be more advanced and personalized in Society D compared to our reality. The potential risks to user data security in the real world were also highlighted as a distinguishing factor. Participants appreciated the attention paid to personal traits, preferences, and the diversity of options available to individuals in Society D. Furthermore, participants acknowledged that our society does have regulations and tests in place for smart health applications, similar to those present in Society D.

In summary, all participants regarded Society D as the most trustworthy society for the development of smart health applications. Its emphasis on choice, customization, respect for diversity, transparency, safety measures, and regulations resonated positively with the participants. From the perspective of the participants, while our society strives towards a similar vision, there are still differences in terms of transparency, artificial intelligence sophistication, and data security. Nonetheless, the existence of regulatory measures and attention to individual preferences and traits were recognized as shared characteristics between our society and Society D.

Society Rankings

Overall, the participants ranked Society D as the most suitable society for the development of smart health applications. Society C followed behind, and Society B was ranked third. Society A received the lowest ranking and was deemed the least suitable for the development of smart health applications.

Figure 4.7 illustrates the rankings given by participants to each society in terms of trustworthiness. It provides a representation of the percentage of participants who ranked each society as their first, second, third, or last choice. Notably, participants had the option to assign the same rank to two societies if they perceived them to have an equal level of trustworthiness. As an example, a participant could have assigned an equal ranking to Society D and Society C, considering them both as their top choices. Consequently, the rankings of the other prototypes were adjusted accordingly, with the remaining societies being ranked third and fourth. No society was given the rank of second in this scenario.

4.5.3 Evaluation Of The Prototypes

The participants successfully completed the predefined tasks for all of the prototypes. Their evaluations varied across the different prototypes. Prototype one was generally perceived as untrustworthy, with only a few participants considering it somewhat trustworthy. On the other hand, all participants found prototype four to be trustworthy. Similarly, almost all participants regarded prototype three as trustworthy, although one participant expressed some doubts and considered it trustworthy to a certain extent. The participants' opinions regarding the trustworthiness of prototype two were more mixed, with three participants finding it untrustworthy, two participants considering it trustworthy, and one participant finding it trustworthy to some degree.

User Perceptions on Prototype 1 (for Society A)

Participants expressed dissatisfaction with the rigidity and lack of flexibility in prototype one. P1 articulated their frustration, stating, "*It is so strict and has limited me to do something that I can't do. I don't like the limitation.*" Similarly, P3 criticized the prototype for its strict scheduling, mentioning, "*It lists the daily activities with the exact time, 7:00 AM, 8:00 AM. I don't like someone else planning my day for me.*" Participants also voiced concerns about the prototype's lack of explanations when notifying them about incomplete activities. Additionally, some participants expressed apprehension about unknown ways the app may monitor them and how it relates to the safety of their data.

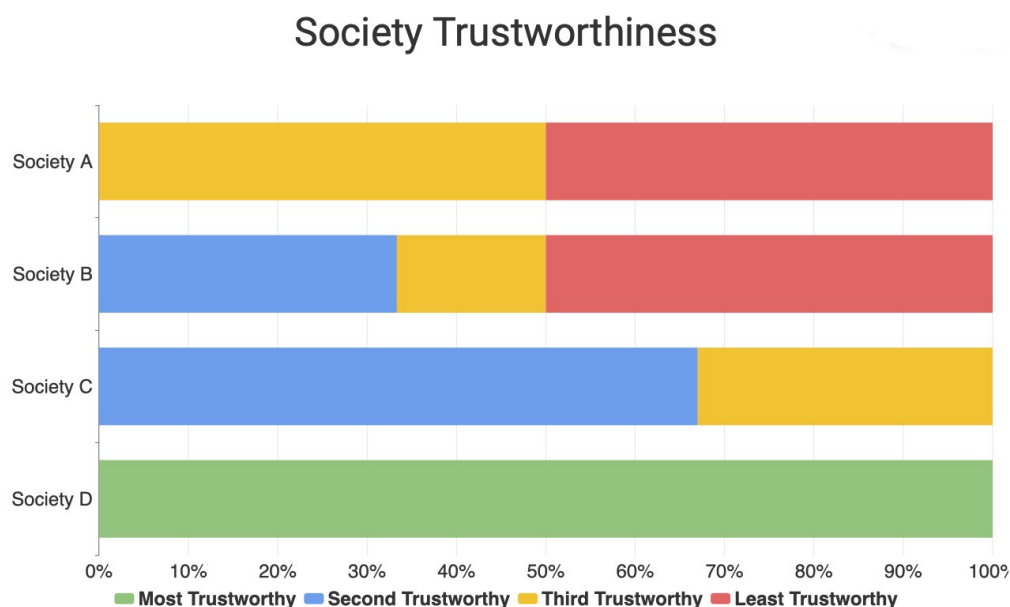


Figure 4.7 The ranking of societies in terms of trustworthiness

The major differences highlighted between this prototype and the applications participants had previously encountered were the lack of choices and freedom, as well as constant monitoring. Some participants struggled to identify any similarities, while others mentioned the existence of a tracker and daily exercise routines, albeit with a notable absence of explanations, as potential areas of overlap.

User Perceptions on Prototype 2 (for Society B)

The majority of participants, again, found prototype two to be overly limiting and restrictive. P6 described it as *“too strict and not flexible.”* Complaining about a similar problem, P5 expressed frustration about being confined to a specific plan without sufficient explanation, stating, *“I don’t understand why I have to walk at this time and cannot walk at any time of the day. It doesn’t give me a reason, it doesn’t tell me why!”* Consequently, they deemed the prototype untrustworthy. However, a minority of participants were satisfied with the safety measures implemented to protect their data, which led them to consider the prototype trustworthy. One participant elaborated on this, stating, *“It also informed me that the data is being monitored solely for my own health, and it won’t be shared with any third parties, which instilled trust in me.”* Additionally, some participants appreciated having a specific

time for their activities, as it relieved them of the burden of planning.

Participants noted that the lack of flexibility was the primary distinction between prototype two and other applications they had encountered. However, they also acknowledged that the prototype offered comprehensive tracking and presented them with daily activities, which served as a notable similarity.

User Perceptions on Prototype 3 (for Society C)

Most participants found prototype three to be trustworthy, despite the fact that it occasionally provided unhealthy diet suggestions upon request. They acknowledged that the explanations and warnings provided were generally sufficient to address their medical safety concerns. However, some participants expressed doubts regarding the safety of data collection methods and the privacy of their information, leading P1 to consider the prototype trustworthy only to a certain extent. This concern was articulated by P1 who remarked, *“It is not clear to me whether this application will share my data with third-party apps or not.”*

On the other hand, participants highlighted the prototype’s individuality and tailored approach as positive aspects, appreciating how it empowered them with choice and agency by explaining the risks and rationale behind each plan and leaving the final decision to them. P5 made the following remark regarding the prototype: *“What I appreciate about this design is that the ultimate choice is still left to the user. You are provided with all the details about why something is good or bad, and then you can choose.”* They further emphasized, *“It mentioned that this option might not be the healthiest, but you are still free to choose it.”*

Participants identified the personalized nature of the prototype and its ability to generate plans *“based on your past data or habits”* (P1) as the primary difference between this prototype and the smart health applications they had previously encountered. The main similarity mentioned by participants was the presence of multiple options and the inclusion of a recipe guide.

User Perceptions on Prototype 4 (for Society D)

Participants expressed high satisfaction with the transparency of prototype four and how it provided them with detailed explanations for each decision. P1 summarized this point by stating, *“The prototype explained the reasoning behind the diagnosis. It also clarified how effective it is and when one should consult a doctor.”* P5 echoed this sentiment, emphasizing that the prototype *“provides the ‘why’ and offers additional details. This level of transparency helps build trust because when users understand what is happening, they are more likely to*

trust the system.” P4 also highlighted the importance of explaining the reasoning behind the final diagnosis, stating that they trust the prototype because “*it compares the data with predefined databases, which gives me confidence in its data accuracy.*” Additionally, participants emphasized the assurance they felt about the safety of their data due to the explanations provided regarding how the data was gathered. P5 particularly emphasized the significance of this “*additional information*” and the ability to edit personal information if needed. The participants also emphasized the significance of having the ability to “*choose*” (P1), “*have freedom*” (P2), and access personalized options, which they considered important factors in establishing trustworthiness.

Regarding similarities, participants identified the ability to enter symptoms and receive a diagnosis as the main commonality between this prototype and the applications they had previous experience with. However, they overwhelmingly pointed to the extensive explanations and transparency as the key differentiating factors that set this prototype apart from regular health apps. P6 also noted that this prototype shows “*The less probable diagnosis and the likelihood so you can see for yourself*” unlike the other application they had an experience with.

Prototype Rankings

Figure 4.7 depicts the trustworthiness rankings assigned by participants to each prototype. It shows the proportion of participants who placed each prototype as their top, second, third, or bottom choice. Notably, participants could rank two prototypes equally if they felt they shared the same trustworthiness level. For instance, if a participant equally ranked Prototype D and Prototype C as top choices, the other prototypes’ rankings were adjusted, resulting in the remaining prototypes ranked as third and fourth. In such cases, no prototype received a second-place ranking.

4.5.4 Match Between Society Descriptions And Prototypes

The findings indicate a reasonable degree of alignment between the prototypes and their corresponding societies, as well as the four criteria of trustworthiness. Specifically, five participants correctly matched Society A with Prototype 1 and Society D with Prototype 4. Similarly, 4 participants correctly associated Society C with Prototype 3. However, only three participants matched Prototype 2 with Society B.

Notably, none of the participants matched Prototypes 3 and 4 which had high levels *explainability* and *fairness* to Society 1 or Society 2 which lacked these characteristics. Similarly, no

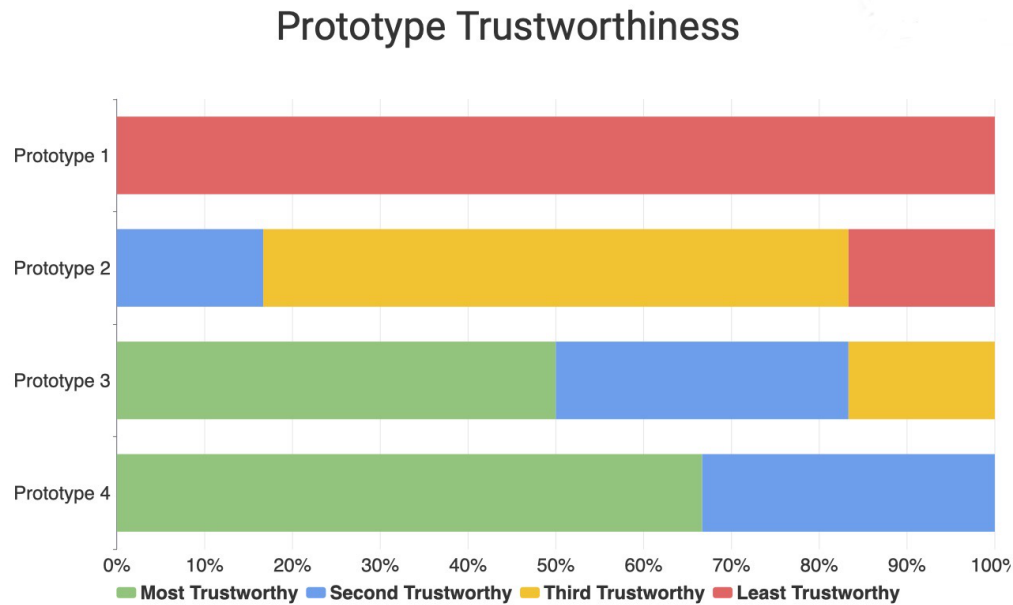


Figure 4.8 The ranking of prototypes in terms of trustworthiness

participant matched Prototype 2 or Prototype 4 which had a high degree of the pairing of "audibility" and "safety" to Society A or Society C which lacked them. This suggests that the prototypes effectively represented the essential characteristics of the societies and the four criteria of trustworthiness.

Overall, these results demonstrate a reasonable alignment between the prototypes and their corresponding societies, indicating that the prototypes adequately reflected the important elements of trustworthiness.

4.6 Discussion

4.6.1 The Importance Of The Four Criteria

The results of our study confirm that the four criteria of *fairness*, *explainability*, *audibility*, and *safety* are essential for the development of trustworthy smart health applications. Prototype D, which exhibited higher levels of all four elements, was consistently considered the highest by the participants in terms of trustworthiness, while prototype A, lacking in these criteria, was consistently considered the worst. The findings also indicate that societies and environments that consider these characteristics essential are perceived as morally

superior and more trustworthy, both for themselves and for the development of smart health applications.

The absence of these four characteristics evoked strong reactions from participants, who described the fictional societies lacking them as dictatorships and military camps, drawing stark comparisons to real-world despotic societies. In contrast, societies that embodied these essential characteristics garnered praise and were viewed as Utopian, emphasizing that these criteria hold intrinsic significance and reflect fundamental aspects of a world. Their presence or absence can offer valuable insights into an environment. Remarkably, even without an explicit framework or guideline, many participants' comments naturally aligned with the four criteria, underscoring their importance and illustrating how they encapsulate ongoing discussions surrounding the trustworthiness of smart health apps. The substantial presence of these four criteria in participants' responses, along with their ability to identify specific elements that align with each criterion, served as a pivotal distinction between the various societies and prototypes. This observation underscored the success of the society descriptions and prototype design in effectively communicating these essential elements and demonstrating the variations in their levels across different societies. It emphasized how these criteria played a central role in shaping participants' perceptions and evaluations.

Furthermore, participants emphasized the significance of trusting the intentions of an application and believing that the team behind it cares about the users. This suggests that the perceived intentions, which may be influenced by the team's reputation, outward appearance, or initial behavior, are important elements contributing to the trustworthiness of smart health applications, alongside the four established criteria. While certain aspects such as the reputation of the company or the general perception of smart applications are beyond the control of designers, further investigation can explore other elements like the visual and functional qualities of an application, which can create a sense of benevolence for users.

In conclusion, our study affirms the importance of the four criteria in the development of trustworthy smart health applications. The existence of these criteria in society can create an optimal environment for the creation of such applications. The findings also highlight the significance of perceived intentions and user trust in the development process.

4.6.2 The Vitality Of Personalizability And Explainability

Personalizability and explainability emerged as essential factors in establishing trustworthiness in smart health applications, according to the participants. The participants emphasized the importance of receiving fair and customized plans and recommendations that catered to their individual needs, characteristics, and preferences. They viewed societies and applica-

tions that followed a one-size-fits-all approach as untrustworthy and lacking in care.

The participants recognized the significance of treating different demographics and subgroups uniquely to foster trust. They valued the ability to customize the application themselves, as it contributed to a sense of fairness and agency. It was observed that some participants were even willing to tolerate certain safety issues and lack of checks and balances if the app provided them with informed choices based on personal preferences, underscoring the value they placed on personal control over health decisions.

Moreover, participants also highlighted the importance of transparency and the ability to understand the rationale behind suggestions or plans. The absence of explanation resulted in an immediate loss of trust, even if safety and medical accuracy were assured. Lack of explainability in societies was associated with undemocratic and manipulative traits.

Interestingly, and similar to fairness, a high level of explainability also sometimes compensated for the absence of safety and audibility. Participants believed that as long as they were aware of the associated risks, they could limit their interactions with low-risk functions or make informed choices regarding the risks involved.

Overall, the findings suggest that the pairing of explainability and fairness is particularly crucial for every day and medically low-risk functions of smart health apps. However, it is important to acknowledge that more sophisticated and sensitive smart health systems, such as advanced diagnostic tools, may have different requirements. Safety becomes especially critical in such systems, as the consequences of inadequate safety measures can be catastrophic.

4.6.3 Giving Users Freedom

The theme of freedom and choice was the most repeated theme among users when explaining their perception of trustworthiness in prototypes and societies for the development of trustworthy applications. While this theme partially overlaps with the fairness criterion discussed earlier, participants' focus on freedom went beyond ensuring the suitability, personalization, and fairness to their individual needs and preferences.

Participants consistently expressed that when an application provides them with the ability to make choices, they are more inclined to trust it and perceive it as trustworthy. This sense of control and mastery over their own journey fosters trust, as they feel respected by the application, which functions as a supportive tool. Participants valued having the final say in their health-related decisions, with the application serving as an aid rather than an authoritative entity.

The emphasis on freedom and choice reflects a fundamental need for user agency and empow-

erment. Participants highlighted that when they have the ability to choose, they perceive a higher level of trust in the application. It signifies a mutual relationship where the application respects their autonomy and empowers them to make decisions. This theme reveals that trustworthiness is not solely determined by functional aspects but also by the extent to which users are empowered through choice.

Moreover, the agency provided to users reassures them that unexpected and potentially risky system functions will not surprise them, as they are guiding the path and making informed decisions. Participants expressed general distrust toward societies that limit user freedom and consider them autocratic. However, they were more willing to accept a degree of control during important health decisions, such as during a public health crisis, as long as clear explanations were provided. This presents a potential conflict between safety and freedom, as trustworthy health applications sometimes necessitate limiting user choices to ensure safety. This potential conflict requires the designers to find a suitable balance that can reassure the users while refraining them from making evidently harmful choices, such balance might prove to be tricky to establish considering the users' s distaste of restrictions on their agency concerning their own health.

Additionally, some participants mentioned a lack of understanding regarding why they should trust an app or a hypothetical government with their health, questioning their intentions. This suggests that trust can be mitigated if the absence of choice comes from a previously trusted source with perceived benevolent intentions, which underscores the importance of perceived intentions discussed earlier.

All in all, the repeated mention of freedom and choice highlights the significance of user agency in establishing trust in applications. When users have the ability to make choices and feel in control of their health journey, their trust in the application increases. The application is viewed as a supportive tool that respects their autonomy and empowers them to make informed decisions.

4.6.4 Design Considerations For The Development Of Trustworthy Applications

Based on the discovered results, the following design considerations can be implemented to enhance the development of smart health applications that are perceived as trustworthy by users:

- **Incorporate the Four Essential Criteria:** The criteria of fairness, explainability, audibility, and safety are crucial for building trust in smart health applications. Ensuring that these criteria are integrated into the design and functionality of the application

can help their perceived level of trustworthiness.

- **Prioritize Personalization and Customization:** Users highly value receiving fair and customized plans and recommendations that cater to their individual needs, characteristics, and preferences. Implementing features that allow users to personalize their experience and make informed choices based on their preferences is thus useful.
- **Emphasize Explainability and Transparency:** Users value transparency and the ability to comprehend the reasoning behind suggestions or plans. Offering clear explanations for the application’s recommendations and decisions can instill confidence and reassure users. Conversely, a lack of explanation can erode trust, even if safety and accuracy are guaranteed.
- **Foster User Agency and Freedom:** Users value the ability to make choices and have control over their health journey. Designing the application in a way that empowers users, respects their autonomy, and gives them the final say in their health-related decisions is essential to create applications that appear trustworthy to users.
- **Establish Trustworthy Intentions:** Users place importance on trusting the intentions of the application and the team behind it. Communicating the benevolent intentions of the development team through reputation, approval from trusted institutions, visual design, and the overall user experience help users trust the intentions behind an application.
- **Balance Safety and Freedom:** It is important to strive to find a suitable balance between ensuring user safety and preserving user freedom. While users value their agency, certain health decisions may require limitations or restrictions to prevent harm. Designing the application in a way that provides clear explanations for these limitations to maintain trust can be an essential but difficult task for the designers.

Following these design considerations, smart health applications can be perceived as more trustworthy by the users and be able to meet users’ needs, foster user trust, and promote positive health outcomes.

4.7 Conclusion

In this phase of the thesis, we adopted a speculative design approach to investigate the importance of four criteria, namely *fairness*, *explainability*, *audibility*, and *safety*, in the development of trustworthy smart health applications. User study participants evaluated

different fictional societies and prototypes created based on these criteria to determine their trustworthiness and reflect on the participants' own notion of trust towards smart health applications.

The results of the study confirmed that the four criteria are indeed crucial for the development of trustworthy smart health applications. Furthermore, personalizability and explainability emerged as important factors for establishing trustworthiness, indicating the value users place on receiving fair and customized services and explanations that cater to their individual needs. The theme of freedom and choice also appeared as an important aspect of trustworthiness. However, designers must find a delicate balance between ensuring user safety and preserving user freedom, as some health decisions may require limitations. The significance of perceived intentions and user trust in the development process of health apps was also discovered as a factor impacting an application's perceived trustworthiness.

Based on these results, we derived design considerations for the development of trustworthy applications. By incorporating these design considerations, smart health applications can be made more trustworthy, meeting users' individual needs, instilling confidence, and promoting positive health outcomes.

CHAPTER 5 CONCLUSION

5.1 Summary of the Work

The two phases of our study aimed to provide valuable insights into users' perceptions toward smart health applications, with a focus on enhancing user trust and improving user experience.

In the first phase, the thematic analysis of user reviews revealed that users' concerns primarily revolved around the accuracy, customizability, and convenience of data input in mobile health applications. The findings underscored the importance of effective communication of the smart component's capabilities and accurate performance to maintain user trust and satisfaction. By addressing these concerns and providing sufficient user control and feedback, designer and developers can better satisfy users' needs in smart health applications.

Building upon the insights from the first phase, the second phase delved deeper into the important criteria influencing the perceived trustworthiness of smart health applications to the users. To provide a less restrictive set of insights into users' experiences and perceptions regarding the trustworthiness of smart health applications, we employed the innovative approach of speculative design. This creative methodology enabled us to explore alternative societal environments for smart health applications and envision more possibilities for the future of smart health applications. By engaging in speculative design, we were able to transcend conventional boundaries and tap into the users' expectations, ideas, and perceptions in a more open-minded and unrestricted manner.

Through the utilization of speculative design, we gained a deeper understanding of the multifaceted nature of user trust in smart health applications. We explored various scenarios, hypothetical societies, and prototypes, allowing us to delve into the nuanced aspects of trustworthiness. This approach facilitated a comprehensive examination of the characteristics and qualities that contribute to user trust, providing valuable insights that can inform the development of trustworthy smart health applications.

The results of the second phase confirmed that these four criteria are indeed crucial for establishing trustworthiness in smart health applications. Prototypes that exhibited higher levels of fairness, explainability, audibility, and safety were consistently ranked higher in terms of trustworthiness by the participants. Conversely, prototypes lacking in these criteria scored lower, indicating their diminished perceived trustworthiness. Participants also emphasized the significance of fair and customized health-related plans suggested to them by smart

health applications, as well as the ability to make informed choices tailored to their individual needs. Personalizability and explainability emerged as essential factors in establishing trustworthiness, highlighting the importance of user empowerment and understanding the rationale behind recommendations and plans. Additionally, the theme of freedom and choice played a central role in participants' perceptions of trustworthiness. Users valued the ability to have control over their health journey and make informed decisions, while still relying on the application as a supportive tool. However, designers must strike a delicate balance between user freedom and ensuring safety, considering the potential conflicts that may arise. This study also underscores the importance of perceived intentions and user trust in the development process. Users place significance on trusting the intentions of the application and the team behind it, as it influences their perception of trustworthiness.

In summary, the combination of both phases of our study has significantly advanced our understanding of trustworthiness in smart health applications. The results have demonstrated the indispensable role of fairness, explainability, audibility, and safety in establishing trust, while the creative insights from speculative design have enriched our understanding of the user's expectations, ideas, and perceptions regarding trust in these applications. Moving forward, these findings can guide the design and development of future smart health applications, fostering user trust and ultimately enhancing the overall user experience.

5.2 Limitations and Future Research

The current study had several limitations that could be addressed in future research. Firstly, the analysis of health applications was limited to a small sample size, which may restrict the generalizability of the findings. Future studies should consider including a larger and more diverse range of applications to provide a more comprehensive understanding of the issues faced by users.

Another limitation was that the reviews analyzed were solely from Android users, potentially overlooking differences in user experiences on other mobile platforms. Future research should aim to include reviews from multiple platforms to capture a broader perspective.

Relying solely on user reviews also limited our understanding of user perceptions, particularly from those who did not leave reviews. Incorporating more exploratory methods, such as interviews or focus groups, can offer deeper insights into the challenges encountered by users of smart health applications.

In the second phase of the study, the participation pool consisted primarily of well-educated individuals, which may have influenced the findings. Future studies should strive for a more

diverse participant sample to account for variations in demographics and their impact on perceptions of trustworthiness.

Additionally, the study's reliance on society descriptions and prototypes, as well as the time constraints imposed in the study design, may have limited the immersive experience for participants. Future research could explore more comprehensive and immersive approaches, such as utilizing multiple design artifacts or interactive simulations, including those in exhibition rooms, to provide a more realistic representation of the scenarios and better stimulate critical discourse.

Moreover, the findings are derived from an artificial scenario or setting, which raises questions about their applicability to real-world situations. Future work should focus on validating these findings in more authentic settings. Furthermore, while it was previously noted that trust tends to be strengthened in risky situations, this aspect was not investigated in the current study. Future research could specifically focus on the role of risk in building trust among users of smart health applications.

Lastly, while this study primarily focused on the issue of trustworthiness in smart health applications, future research should explore other important factors such as addressing inaccuracies or mitigating the disproportionate emphasis on monetary gains by the applications. Taking a holistic approach to investigating these issues will contribute to the development of user-friendly and effective smart health applications.

By addressing these limitations and conducting future research in these areas, we can further advance our understanding of the challenges faced by users of smart health apps and contribute to the development of improved and trustworthy applications.

Lastly, the number of participants was relatively small, and involving a larger pool of participants could have yielded additional insights and increased clarity in our findings.

REFERENCES

- [1] E. Toreini *et al.*, “The relationship between trust in ai and trustworthy machine learning technologies,” in *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency*, ser. FAT* ’20. New York, NY, USA: Association for Computing Machinery, 2020. [Online]. Available: <https://doi.org/10.1145/3351095.3372834>
- [2] P. Bhatt *et al.*, “Emerging artificial intelligence-empowered mhealth: Scoping review,” *JMIR Mhealth Uhealth*, vol. 10, no. 6, p. e35053, Jun 2022.
- [3] I. Vaghefi and B. Tulu, “The continued use of mobile health apps: Insights from a longitudinal study,” *JMIR Mhealth Uhealth*, vol. 7, no. 8, p. e12983, Aug 2019. [Online]. Available: <http://mhealth.jmir.org/2019/8/e12983/>
- [4] W. Peng *et al.*, “A qualitative study of user perceptions of mobile health apps,” *BMC Public Health*, vol. 16, p. 1158, 2016.
- [5] M. Patel, D. Asch, and K. Volpp, “Wearable devices as facilitators, not drivers, of health behavior change,” *JAMA*, vol. 313, no. 5, pp. 459–460, 2015.
- [6] E. Topol, “High-performance medicine: the convergence of human and artificial intelligence,” *Nat Med*, vol. 25, no. 1, pp. 44–56, 2019.
- [7] S. Deterding *et al.*, “From game design elements to gamefulness: Defining gamification,” in *Proceedings of the 15th International Academic MindTrek Conference: Envisioning Future Media Environments, MindTrek 2011*, vol. 11, 09 2011, pp. 9–15.
- [8] J. Adler-Milstein *et al.*, “Electronic health record adoption in us hospitals: Progress continues, but challenges persist,” *Health Aff (Millwood)*, vol. 34, no. 12, pp. 2174–2180, Dec 2015.
- [9] Z. Obermeyer and E. Emanuel, “Predicting the future - big data, machine learning, and clinical medicine,” *N Engl J Med*, vol. 375, no. 13, pp. 1216–1219, Sep 29 2016.
- [10] H. Thimbleby, “Technology and the future of healthcare,” *J Public Health Res*, vol. 2, no. 3, p. e28, Dec 1 2013.
- [11] A. Dunne and F. Raby, *Speculative Everything: Design, Fiction, and Social Dreaming*. The MIT Press, 2013.

- [12] J. Niess *et al.*, “‘i don’t need a goal’: Attitudes and practices in fitness tracking beyond weird user groups,” in *Proceedings of the 23rd International Conference on Mobile Human-Computer Interaction*, ser. MobileHCI ’21. New York, NY, USA: Association for Computing Machinery, 2021. [Online]. Available: <https://doi.org/10.1145/3447526.3472062>
- [13] N. Lau *et al.*, “Android and iphone mobile apps for psychosocial wellness and stress management: Systematic search in app stores and literature review,” *JMIR Mhealth Uhealth*, vol. 8, no. 5, p. e17798, May 2020. [Online]. Available: <http://mhealth.jmir.org/2020/5/e17798/>
- [14] C. Kao and D. Liebovitz, “Consumer mobile health apps: Current state, barriers, and future directions,” *PM R*, vol. 9, no. 5S, pp. S106–S115, May 2017.
- [15] G. Dove *et al.*, “Ux design innovation: Challenges for working with machine learning as a design material,” in *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, ser. CHI ’17. New York, NY, USA: Association for Computing Machinery, 2017, p. 278–288. [Online]. Available: <https://doi.org/10.1145/3025453.3025739>
- [16] S. Amershi *et al.*, “Guidelines for human-ai interaction,” in *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, ser. CHI ’19. New York, NY, USA: Association for Computing Machinery, 2019, p. 1–13. [Online]. Available: <https://doi.org/10.1145/3290605.3300233>
- [17] S. T. Völkel *et al.*, “What is ‘intelligent’ in intelligent user interfaces? a meta-analysis of 25 years of iui,” in *Proceedings of the 25th International Conference on Intelligent User Interfaces*, ser. IUI ’20. New York, NY, USA: Association for Computing Machinery, 2020, p. 477–487. [Online]. Available: <https://doi.org/10.1145/3377325.3377500>
- [18] R. Vasa *et al.*, “A preliminary analysis of mobile app user reviews,” in *Proceedings of the 24th Australian Computer-Human Interaction Conference, OzCHI 2012*, 11 2012.
- [19] A. K. Ghosh *et al.*, “Safety vs. surveillance: What children have to say about mobile apps for parental control,” in *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, ser. CHI ’18. New York, NY, USA: Association for Computing Machinery, 2018, p. 1–14. [Online]. Available: <https://doi.org/10.1145/3173574.3173698>
- [20] T. Koulouri, R. D. Macredie, and D. Olakitan, “Chatbots to support young adults’ mental health: An exploratory study of acceptability,” *ACM Trans. Interact. Intell. Syst.*, vol. 12, no. 2, jul 2022. [Online]. Available: <https://doi.org/10.1145/3485874>

- [21] C. Phillips *et al.*, “Identifying commercial games with therapeutic potential through a content analysis of steam reviews,” *Proceedings of the ACM on Human-Computer Interaction*, vol. 5, 10 2021.
- [22] J. A. Simpson, “Psychological foundations of trust,” *Current Directions in Psychological Science*, vol. 16, no. 5, pp. 264–268, 2007. [Online]. Available: <https://doi.org/10.1111/j.1467-8721.2007.00517.x>
- [23] N. Ashraf, I. Bohnet, and N. Piankov, “Decomposing trust and trustworthiness,” *Exp Econ*, vol. 9, pp. 193–208, 2006.
- [24] K. Siau and W. Wang, “Building trust in artificial intelligence, machine learning, and robotics,” *Cutter Business Technology Journal*, vol. 31, pp. 47–53, 03 2018.
- [25] J. Tripp, “Technology, humanness and trust: Rethinking trust in technology.” *Journal of the Association for Information Systems*, vol. 16, p. 1, 12 2015.
- [26] P. Wicks and E. Chiauzzi, “‘trust but verify’—five approaches to ensure safe medical apps,” *BMC Med*, vol. 13, p. 205, Sep 25 2015.
- [27] U. Albrecht, “Transparency of health-apps for trust and decision making,” *J Med Internet Res*, vol. 15, no. 12, p. e277, Dec 30 2013.
- [28] J. Torous *et al.*, “Clinical review of user engagement with mental health smartphone apps: evidence, theory and improvements,” *BMJ Ment Health*, vol. 21, no. 3, pp. 116–119, 2018. [Online]. Available: <https://mentalhealth.bmj.com/content/21/3/116>
- [29] J. Auger, “Speculative design: Crafting the speculation,” *Digital Creativity*, vol. 24, 03 2013.
- [30] T. Hoang *et al.*, “What can speculative design teach us about designing for healthcare services?” in *Proceedings of the 30th Australian Conference on Computer-Human Interaction*, ser. OzCHI ’18. New York, NY, USA: Association for Computing Machinery, 2018, p. 463–472. [Online]. Available: <https://doi.org/10.1145/3292147.3292160>
- [31] R. LC and D. Mizuno, “Designing for narrative influence: Speculative storytelling for social good in times of public health and climate crises,” in *Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems*, ser. CHI EA ’21. New York, NY, USA: Association for Computing Machinery, 2021. [Online]. Available: <https://doi.org/10.1145/3411763.3450373>

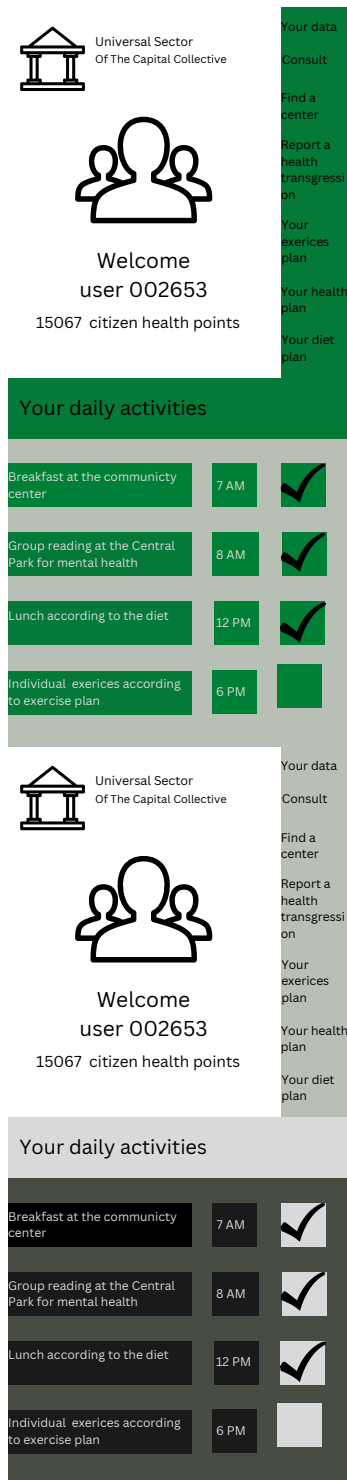
- [32] J. Smith, G. Klumbyte, and R. L. Britton, "From 'if-then' to 'what if?' rethinking healthcare algorithmics with posthuman speculative ethics," *Nursing Philosophy*, vol. 24, no. 3, p. e12447, 2023. [Online]. Available: <https://onlinelibrary.wiley.com/doi/abs/10.1111/nup.12447>
- [33] M. Hebrok and H. Mainsah, "Skinny as a bird: Design fiction as a vehicle for reflecting on food futures," *Futures*, vol. 141, p. 102983, 2022. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0016328722000817>
- [34] T. Harwood, T. Garry, and R. Belk, "Design fiction diegetic prototyping: a research framework for visualizing service innovations," *Journal of Services Marketing*, vol. ahead-of-print, 12 2019.
- [35] E. Tseklevs *et al.*, "Rethinking how healthcare is conceptualised and delivered through speculative design in the uk and malaysia: A comparative study," *The Design Journal*, vol. 22, no. sup1, pp. 429–444, 2019. [Online]. Available: <https://doi.org/10.1080/14606925.2019.1595430>
- [36] A. Mustafa *et al.*, "User engagement and abandonment of mhealth: A cross-sectional survey," *Healthcare (Basel)*, vol. 10, no. 2, p. 221, Jan 24 2022.
- [37] A. Strzelecki, "Application of developers' and users' dependent factors in app store optimization," *International Journal of Interactive Mobile Technologies (iJIM)*, vol. 14, pp. 91–106, 08 2020.
- [38] A. Jacovi *et al.*, "Formalizing trust in artificial intelligence: Prerequisites, causes and goals of human trust in ai," in *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, ser. FAccT '21. New York, NY, USA: Association for Computing Machinery, 2021, p. 624–635. [Online]. Available: <https://doi.org/10.1145/3442188.3445923>
- [39] D. Johnson and K. Grayson, "Cognitive and affective trust in service relationships," *Journal of Business Research*, vol. 58, no. 4, pp. 500–507, 2005, special Section: Attitude and Affect. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0148296303001401>
- [40] R. Borum, "The science of interpersonal trust," *Randy Borum*, 01 2010.
- [41] J. A. Simpson, "Foundations of interpersonal trust," *Social psychology: Handbook of basic principles*, vol. 2, pp. 587–607, 2007.

- [42] M. Malpass, *Critical Design in Context: History, Theory, and Practice*. Bloomsbury Academic, 2019. [Online]. Available: <https://books.google.ca/books?id=2qCgxgEACAAJ>
- [43] D. MacKenzie and J. Wajcman, Eds., *The Social Shaping of Technology*. Buckingham, UK: Open University, 1999.
- [44] P. Braveman and L. Gottlieb, “The social determinants of health: It’s time to consider the causes of the causes,” *Public Health Reports*, vol. 129, no. 1_suppl2, pp. 19–31, 2014, pMID: 24385661. [Online]. Available: <https://doi.org/10.1177/00333549141291S206>

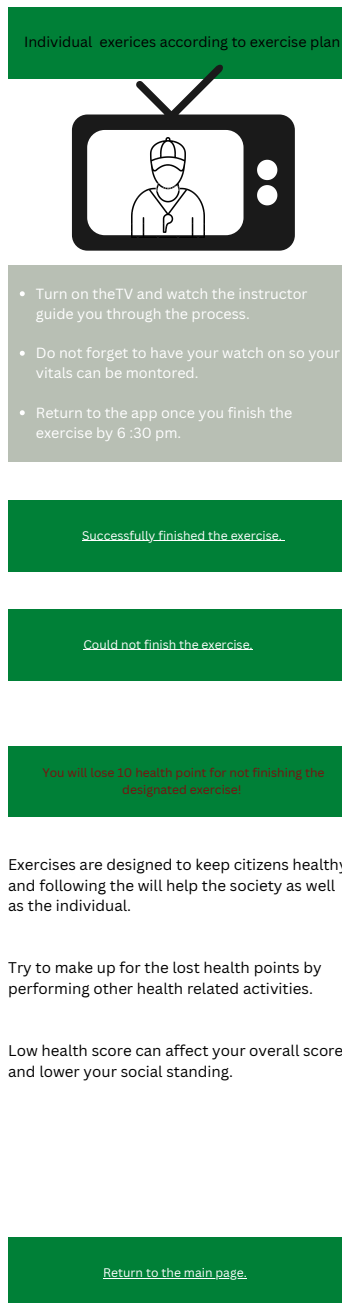
APPENDIX A THE PROTOTYPES

Here are the four prototypes used in the second phase of the study:

Prototype 1



Prototype 1



Prototype 1

Thank you for finishing the designated exercise.

Our data collected during the duration of the exercise shows that you unfortunately were not able to perform the exercise plan correctly. This result in you gaining no extra health points.

learning if they have not performed an activity correctly is for the benefit of the citizens themselves as it allows them to learn and become netter.

Tenacity is an important virtue so try again next time.

[Return to the main page.](#)



Universal Sector
Of The Capital Collective



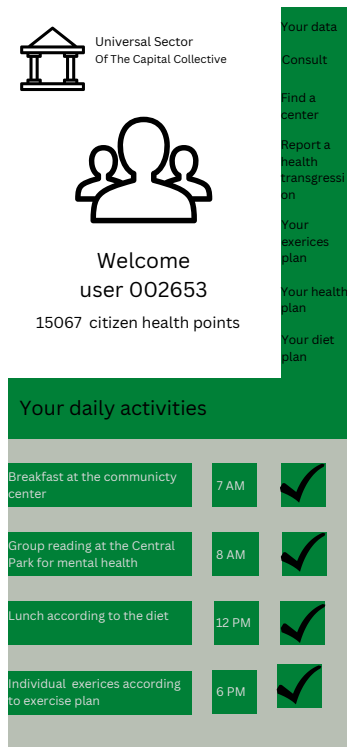
Welcome
user 002652
15057 citizen health points

Your data
Consult
Find a
center
Report a
health
transgression
Your
exercises
plan
Your health
plan
Your diet
plan

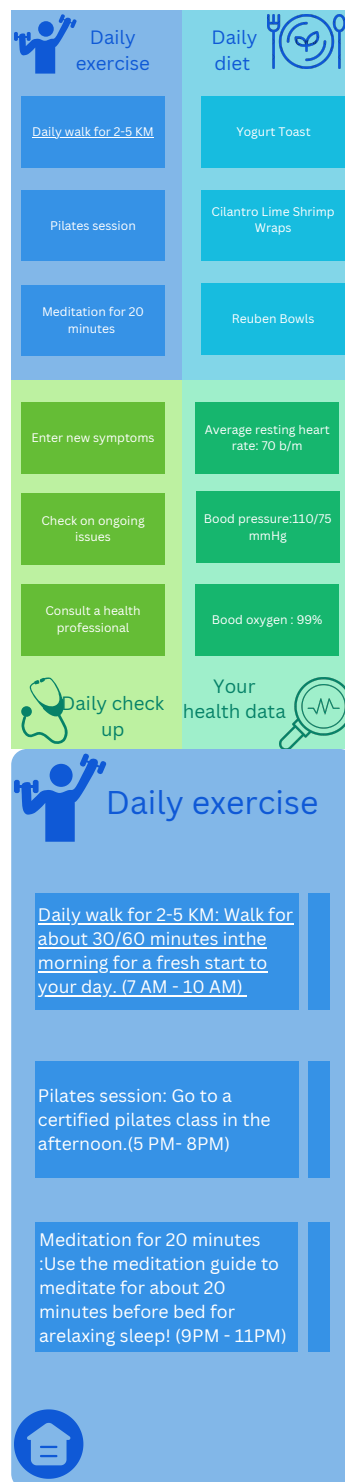
Your daily activities

Breakfast at the community center	7 AM	✓
Group reading at the Central Park for mental health	8 AM	✓
Lunch according to the diet	12 PM	✓
Individual exercises according to exercise plan	6 PM	✗

Prototype 1



Prototype 2



Prototype 2



Daily walk

Daily walk for 2-5 KM: Walk for about 30/60 minutes in the morning for a fresh start to your day. (7 AM - 10 AM)

Mark the start time and location of your walk, for a safer walk wear your health rings on three of your fingers so that your vitals can be monitored at all times.

The effectivity and safety of this exercise plan has been tested and proven by health experts. Your data will be gathered by the government and shared with no other third party.

[Start my walk!](#)





Daily walk

Reminder: Your data will be confidential and protected.

Time 12 PM

Location Use my current location

[Start my walk!](#)



Prototype 2



Daily walk

Time

11 PM

For the best results the starting Tim should be between 7 AM to 10AM

Location

2500 Chem. de Polytechnique, Montréal

Start my walk!





Daily walk

Time

9:45 AM

Location

2500 Chem. de Polytechnique, Montréal

Start my walk!



Prototype 2



Daily walk

Time	9:45 AM	10:15 AM
Starting location	2500 Chem. de Polytechnique, Montréal	
Current location	Snowdon, Côte-Des-Neiges—Notre-Dame-De-Grâce, Montreal, QC	
Distance walked	2.8 KM	

Finish my walk!





Daily walk

Time	9:45 AM	10:15 AM
Starting location	2500 Chem. de Polytechnique, Montréal	
End location	Snowdon, Côte-Des-Neiges—Notre-Dame-De-Grâce, Montreal, QC	
Distance walked	2.8 KM	



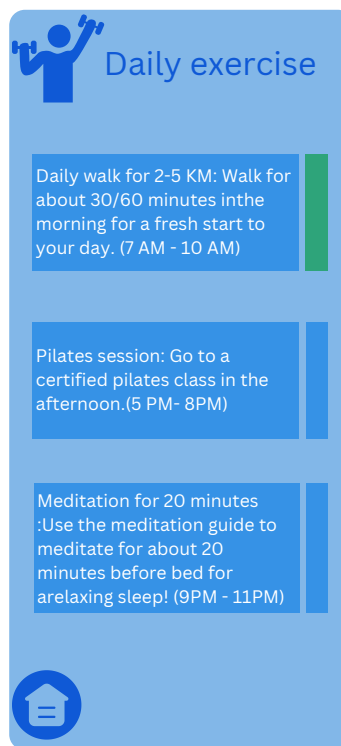
You successfully finished your daily walk!



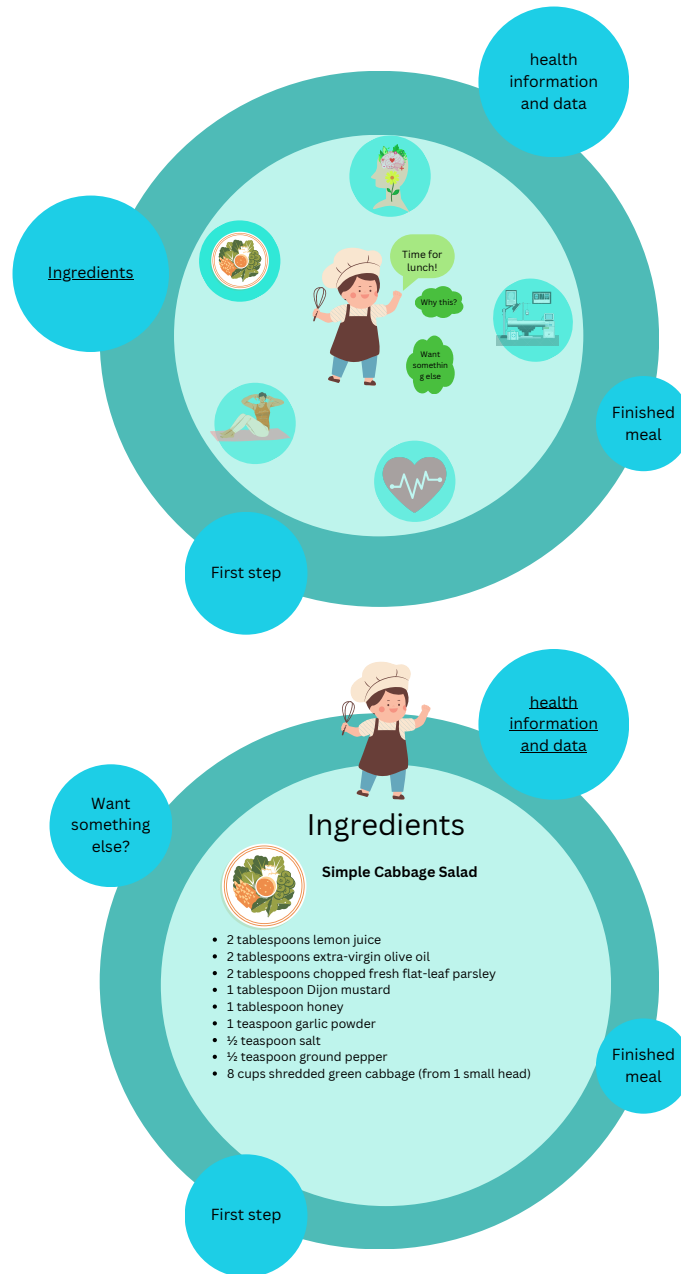
Back to the daily exercises



Prototype 2



Prototype 3



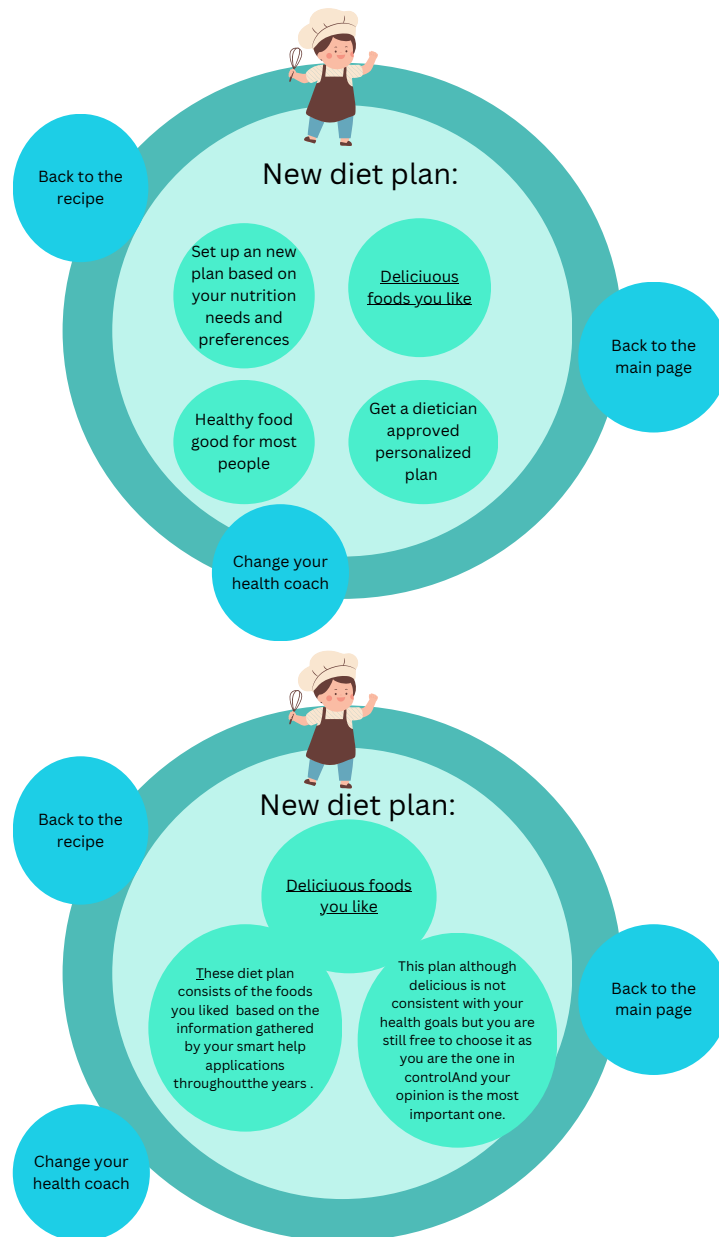
Prototype 3



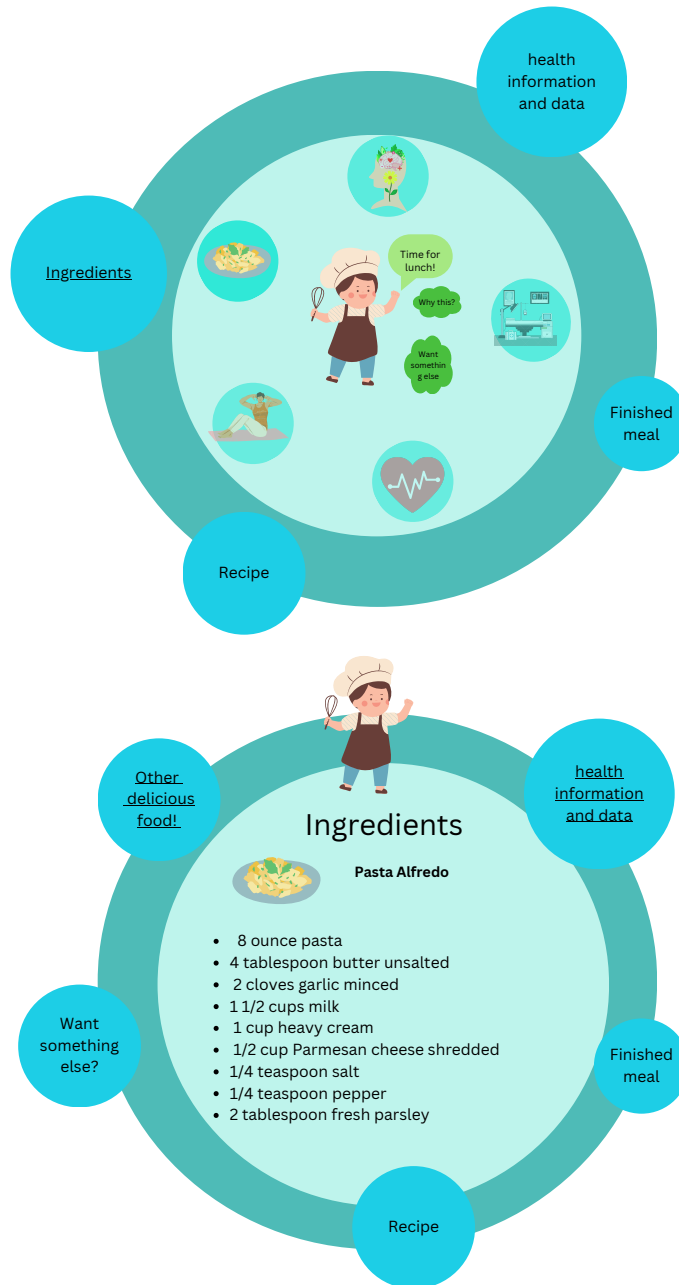
Prototype 3



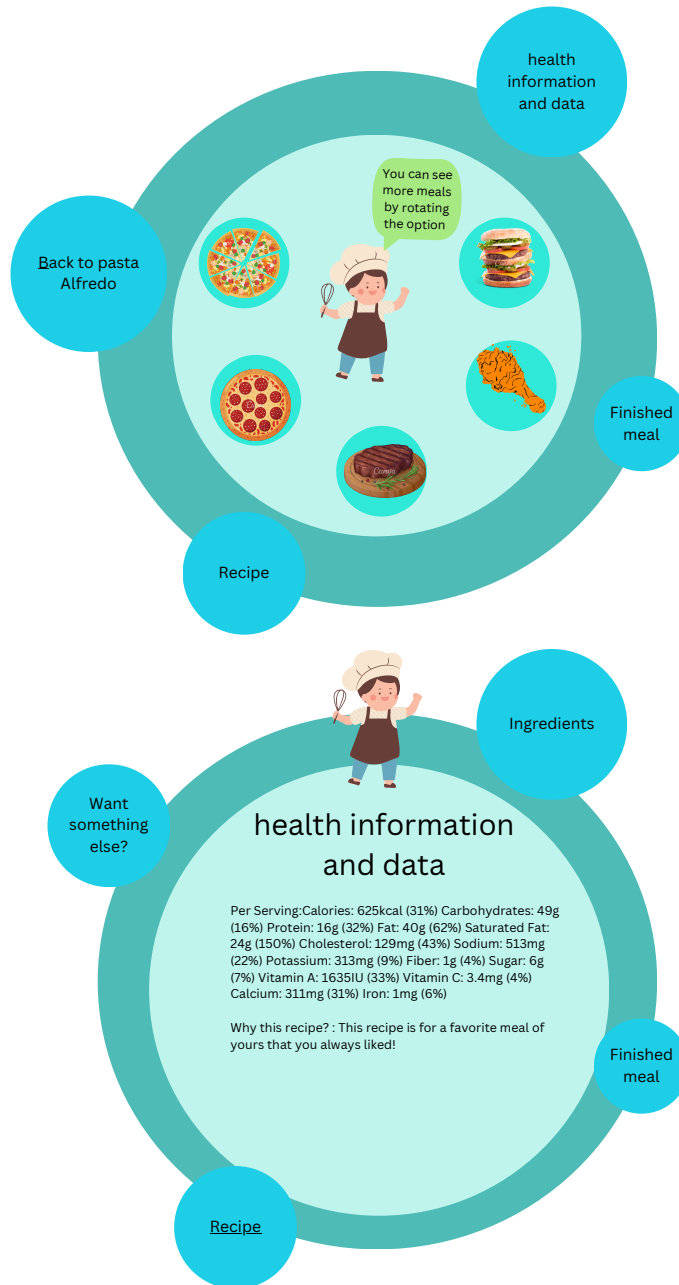
Prototype 3



Prototype 3



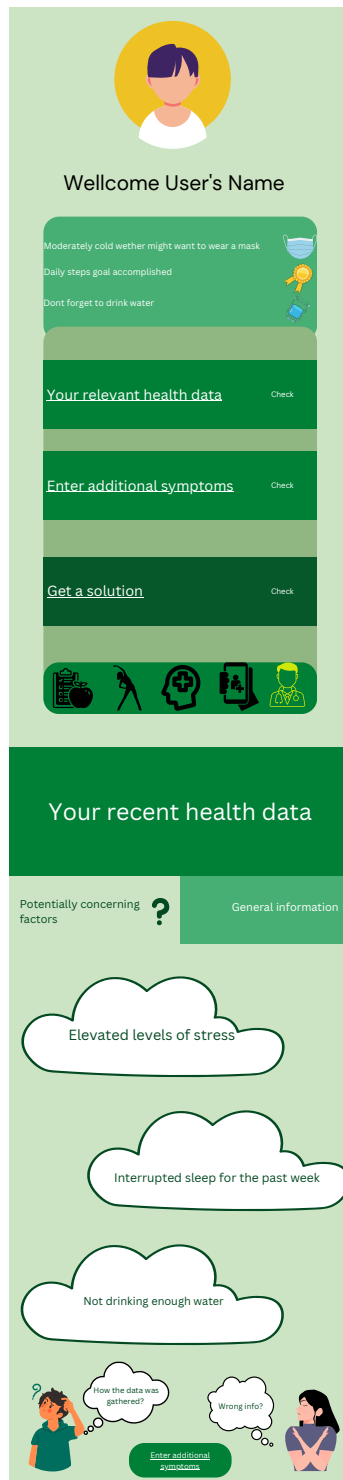
Prototype 3



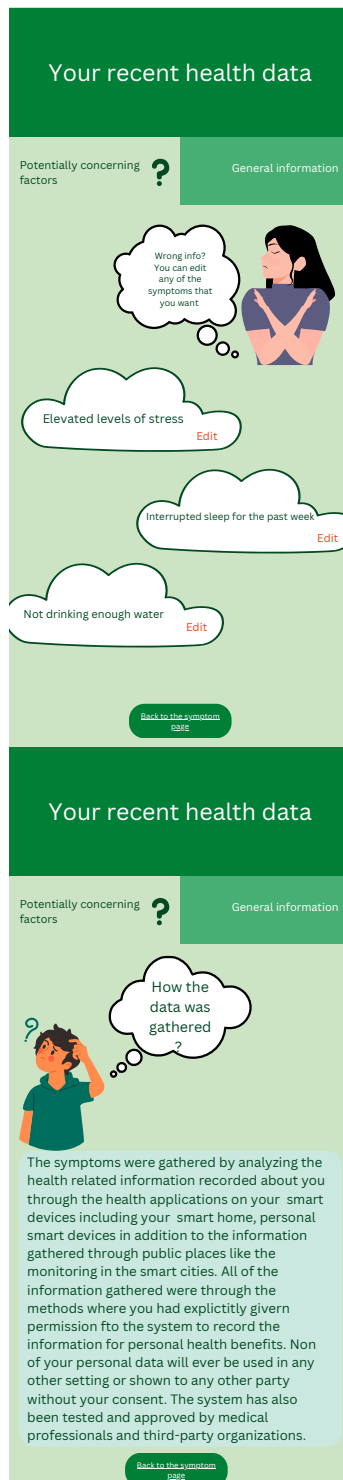
Prototype 3



Prototype 4



Prototype 4



Prototype 4

Enter additional symptoms

choose from the list by typing or point to the to the area of the body to get a limited list

Head ...

headache
Head injury
lightheadedness





Added symptoms

Enter additional symptoms

choose from the list by typing or point to the to the area of the body to get a limited list

...





Added symptoms

1) Headache

Finished entering symptoms

Prototype 4

Enter additional symptoms

Describe your symptoms in detail



Added symptoms

1) Headache


How long it has been going on?

2 days

Severity from one to ten?

4

When does it usually happen?

In the morning

Additional information

[Finished with this symptom](#)

Get a diagnosis



Why this diagnosis?

Your diagnosis based on your data and symptoms

Tension headache

How effective is this?

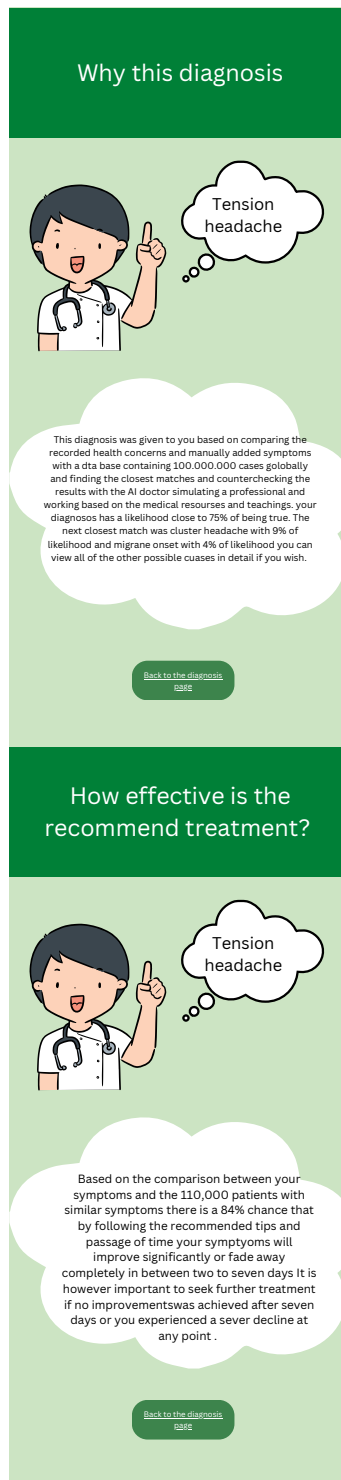
Treatment

1. Over the counter pain medication
2. Stress management methods like meditation.
3. Hydrating more.
4. Sleeping earlier each night.

Important recommendation

If you continue to feel the pain after a week or if the pain gets worse or additional symptoms follow, play a visit to your doctor or go to a hospital depending on the severity of the problem.

Prototype 4



APPENDIX B THE TEST SCRIPT FOR THE USER STUDY

Here is the test script used for the user study sessions:

Thank you very much for participating in this research study. As you have read in the consent form, we are focused on investigating design principles and guidelines that address user trust issues in data-driven systems. In this study, we particularly focus on smart health apps, such as diet apps, fitness apps, and diagnostic tools. We will ask you questions regarding your general perspectives and experiences related to trust in those applications and the effects of societal norms on them. We will also ask you to review several design ideas and provide comments on them.

Before we get started, I would like to remind you that this study will be video recorded. We will strive to keep the confidentiality about your identity and the data. You have the right to end your participation at any point. I would also like to remind you that there is no right or wrong way to complete the tasks or answer the questions. We are not testing your ability to complete the tasks; instead, we want to understand how you answer the question and approach the tasks. Do you have any questions before we proceed? [Answer participants' questions.]

General Perspectives and Experiences

First, we have several questions to ask about your general perspectives and experiences regarding trust in "smart" health applications: [Ask follow-up questions when interesting answers emerge]

- What kind of health applications have you used?
- What kind of "smart" components are you aware of in those applications?
- What is your general experience using those health applications?
- What is your overall point of view regarding the trustworthiness of "smart" health applications?
- Please tell us about a positive situation or experience in which you fully trust a "smart" health application. [Follow-up on factors that led to trust.]
- Please tell us about a negative situation or experience in which you do not trust a "smart" health application. [Follow-up on factors that led to distrust.]

- What kind of broader societal issues do you think have contributed to your experience discussed before? [Enumerate each experience to remind the participants]
- In your opinion, what would be an ideal society that can facilitate more trustworthy technologies and applications?

Demographic Questions

- What gender do you identify as?
- What is your age or age range?
- What is the highest degree or level of education you have completed?
- Where do you currently reside?

Reviewing Society Descriptions

Now we will ask you to read the society description. After reading, please notify us so we can ask you the relevant questions.

[Ask the following questions after the participants are done reading the description for society A :]

- What do you think about this society?
- To what degree do you think the designs reflected the conditions of the society? What elements of the designs lead you to this conclusion?
- Do you think the situation in this society is optimal for the development of trustworthy applications?
- What do you think is the biggest similarity between the characteristics of this society and the characteristics of our current society?
- What do you think is the biggest difference between the characteristics of this society and the characteristics of our current society?

[Move on to the next society until all of the society descriptions are finished]

[Ask the following questions after all societies are reviewed.]

- Please rank the societies in terms of trustworthiness, you can rank two societies at the same level of trustworthiness

Reviewing Corresponding Prototypes

Next, we will show you four preliminary designs for health-related applications corresponding to the four hypothetical societies you just read about. For each preliminary design, we will ask you to complete a number of tasks. After reviewing each design, we will ask you to read a description of that hypothetical society.

[Randomly select a design after the tasks and questions are finished, move on to another design, and continue like that until all of the designs for society A, B, C, and D are finished.]

[Hand the design for society A to the participant]

Please complete the following tasks:

- Take a look at the first page of the application
- Examine today's activity section
- Go to the earliest unfinished activity and try to complete it based on the given interaction
- When you arrive at the page with the picture of TV, look at the TV and watch the exercise
- Based on whether you think you are able to finish the exercise or not, select "successfully finished the exercise" or "could not finish the exercise"
- Read the results and try to see if you can find an explanation for them
- Go back to the home page

After reviewing the design, ask the following questions:

- What are the two things you like about the preliminary design?
- What are the two things you dislike about the preliminary design?
- How important do you think the application represented in this design needs to be trustworthy? Please explain your reasoning.
- How trustworthy do you find this design is? Please explain your reasoning.

- What kind of societal issues, such as social norms and rules, do you think will lead to this type of design? [Ask for specifics.]
- Thinking about the experience you just had interacting with this design, what do you think is the biggest similarity between this experience that you just had and the experiences you have had before interacting with other apps in your daily life?
- Thinking about the experience you just had interacting with this design, what do you think is the biggest difference between this experience that you just had and the experiences you have had before interacting with other apps in your daily life?

[Hand the design for society B to the participant]

Please complete the following tasks:

- Take a look at the first page of the application
- Go to the daily exercise section and select the earliest recommended exercise
- Put on the health check rings
- Try to start the walk by adding your current location at 12 pm
- Look to see if there is additional information available about how you can select this time.
- Retry to start the walk at 9:45 am
- Finish the walk and return to the main exercise page

[Ask the questions mentioned above and read the description of society B to the participant]

[Hand the design for society C to the participant]

Please complete the following tasks:

- Select the recommended lunch icon
- Look at the ingredients
- Look at the health data
- Look at the recipe and go to the finished meal
- Try to get a new recipe by selecting "I want something else"

- Decide to set up a new plan
- Select the delicious food option instead of healthier options
- Go to the main page and select the new recipe by clicking on the food item
- Look at the ingredients, then health information, then the recipe
- Go to the finished product page

[Ask the questions mentioned above and read the description of society C to the participant]

[Hand the design for society D to the participant]

Please complete the following tasks:

- Take a look at the first page of the application
- Go to the diagnostic section by clicking on the doctor icon
- See the current symptoms and check how they were gathered by looking at the relevant health data
- See if there is an option for changing them (no need to actually change them)
- Go back to the symptom page and start adding new symptoms by adding "headache" by selecting it and then press "finished entering symptoms"
- Add 2 days for duration and 4 for the severity and mornings for the time for the headache from the symptoms list for headache
- Go to the diagnostic section by clicking on "finished with this symptom"
- Read the diagnostic given to you and the recommended treatment and important recommendations
- Check why this diagnostic was given
- Go back to the diagnostic page and check how effective the treatment is

[Ask the questions mentioned above and read the description of society D to the participant]

[Ask the following questions after all designs are reviewed.]

- Please rank the designs in terms of trustworthiness, you can rank two designs at the same level of trustworthiness.

- Please provide your mapping of the societies to the corresponding designs at your convenience within the next week. You will receive a link to access the descriptions and designs for reference during the process.

Please share if you have any questions, concerns or suggestions.

Thank you so much for your participation!

APPENDIX C THE CODE BOOK

Below is a section of the code book derived from the analysis of users' reviews during the first phase of the study on smart health applications:

ATLAS.ti Report

playstore_app_reviews

Codes

Report created by Banafsheh Mohajeri on Jul 21, 2023

• bad deisgn (related to UI or UX): bland design

Quotations:

8:10 ¶14, Also it looks very plain in WebMD

Groups:

Bad design (related to UI and UX)

• bad deisgn (related to UI or UX): incomprehensible design and information

Quotations:

8:4 ¶12, Allergy pollen chart is not readable. in WebMD
8:7 ¶13, IT IS NOT COLOR CODED ANYMOR in WebMD
8:8 ¶14, The allergy screen is partially not visible in WebMD
8:11 ¶14, needs the different colors for the intensity of the pollutants and the ma in WebMD

Groups:

Bad design (related to UI and UX)

• bad deisgn (related to UI or UX): notifications not aggressive enough for the target users

Quotations:

8:25 ¶13, reminder notifications are not aggressive enough. People with cognitive issues such as Alzheimer's o... in WebMD

Groups:

Bad design (related to UI and UX)

• Bad design (related to UI and UX) :too many functions

Quotations:

8:14 ¶16, I want a simple widget to show me what I need to know. Stop shoving a plethora of useless 'info' in... in WebMD
8:15 ¶13, -BRING BACK THE ORIGINAL, BY ITSELF, ALLERGY WIDGET!! YOU MOVED IT TO BUNDLE WITH THIS APP in WebMD
8:32 ¶16, I also do not like all the added features, I have separate apps for those and it's creating clutter... in WebMD

Groups:

Bad design (related to UI and UX)

• Bad design (related to UI and UX): hard to understand

Quotations:

6:5 ¶4, I tried to cancel subscription after trial and cannot figure out how! in YAZIO Calorie Counter & Intermittent Fasting App
8:34 ¶12, I don't know specifically where it's coming from in WebMD

Groups:

Bad design (related to UI and UX)

• Bad design (related to UI and UX): unclear serving size

Quotations:

6:2 ¶12, I do have a problem with how unclear serving sizes can be. in YAZIO Calorie Counter & Intermittent Fasting App

Groups:

Bad design (related to UI and UX)

• Bad design (related to UI and UX) :hard to use

Quotations:

1:56 ¶18, So I have to open it every time I start walking anywhere, open it several times in the process, open... in Google Fit
1:59 ¶19, The option to always switch from heart points to steps is a little less user friendly in Google Fit
6:20 ¶12, The app needs to have an easier way to develop verified foods with cups, teaspoons, and Tablespoons... in YAZIO Calorie Counter & Intermittent Fasting App
6:22 ¶13, Too cumbersome to actually track food every day. in YAZIO Calorie Counter & Intermittent Fasting App
6:35 ¶19, the way to add food information is a bit complicated in YAZIO Calorie Counter & Intermittent Fasting App
8:2 ¶1, To put in Mom's birthday, I was forced to use calendar (5+ minutes to scroll back to 1930) in WebMD
8:23 ¶12, If only the app did, too. This app makes you type in what specific organ/part is causing pain in WebMD
8:29 ¶15, Now you must enter it all manually, and there are no follow up questions in WebMD
8:31 ¶16, but now I don't have a place to start. in WebMD

Groups:

Bad design (related to UI and UX)

• Bad design (related to UI and UX) :manual and time consuming method of adding info

Quotations:

1:56 ¶18, So I have to open it every time I start walking anywhere, open it several times in the process, open... in Google Fit
1:63 ¶11, I need to manually set track my activities which is not convenient in Google Fit
1:66 ¶13, I was manually adding the data for some time but what's the point in app when I have to add the dat... in Google Fit
1:77 ¶10, Every day I have to manually adjust Fit step/activity info as it is off by the thousands in my Diabe... in Google Fit
8:33 ¶12, app makes you type in what specific organ/part is causing pain in WebMD
8:38 ¶18, I take various medications and have it listed on the reminder section, however you have to select ea... in WebMD

Groups:

Bad design (related to UI and UX)

• Bad design(Update Problems): lost functions after update or new version

Quotations:

1:40 ¶2, there is obviously an issue with one of the recent updates in Google Fit
1:50 ¶7, Every time Google fit updates the app stops working on my Suunto 7 watch, and on my phone in Google Fit
1:60 ¶11, after last update it is not at all working in Google Fit
8:22 ¶12, The old online symptom checker would let you select an area on the body where you had pain. If only... in WebMD
8:26 ¶14, I miss the old symptom checker in WebMD
8:28 ¶15, this app used to be great in WebMD
8:30 ¶16, I miss the old symptom checker interface in WebMD

Groups:

Bad design (update problems)

• Bad design(Update Problems): Would not work without updating

Quotations:

7:10 ¶5, if there's an update, it won't work in Step Tracker - Pedometer

Groups:

Bad design (update problems)

• Good deisgn (related to UI and UX): Good visuals and look

Quotations:

1:57 ¶19, interface is good in Google Fit
5:13 ¶16, with such soothing colors! in Headspace
5:14 ¶15, Love the lessons along with the visualizations the best in Headspace
6:6 ¶15, Much nicer looking than MyFitnessPal in YAZIO Calorie Counter & Intermittent Fasting App
6:26 ¶17, Nice interface in YAZIO Calorie Counter & Intermittent Fasting App

Groups:

Good design (related to UI and UX)

• Good design (effective feedback)

Quotations:

5:40 ¶17, So I emailed the customer service on the email address given here in the play store and got a reply... in Headspace

Groups:

good design (other aspects)

- **Good design (other aspects): being challenging**

Quotations:

- 🕒 3:40 ¶20, this fitness app is Really challenging in Six Pack in 30 Days

Groups:

- 👤 good design (other aspects)
-

- **Good design (other aspects): reasonable amount of ads**

Quotations:

- 🕒 3:8 ¶5, it doesn't obstruct the use of the app too much in Six Pack in 30 Days
- 🕒 3:18 ¶2, not to many ads in Six Pack in 30 Days
- 🕒 3:20 ¶4, no ads while workout ads after workout that ia the best in Six Pack in 30 Days
- 🕒 3:28 ¶9, Not too many adds as wel in Six Pack in 30 Days
- 🕒 3:32 ¶14, The ads are fairly non-obnoxious in Six Pack in 30 Days

Groups:

- 👤 good design (other aspects)
-

- **Good design (other aspects): use minimal data**

Quotations:

- 🕒 7:2 ¶1, uses minimal data while running in the background in Step Tracker - Pedometer

Groups:

- 👤 good design (other aspects)
-

- **Good design (other aspects): works offline**

Quotations:

- 🕒 3:22 ¶4, yoy can work offline if u want in Six Pack in 30 Days

Groups:

- 👤 good design (other aspects)
-

- **Good design (related to UI and UX): Easy to use**

Quotations:

- 🕒 3:16 ¶2, easy to follow in Six Pack in 30 Days
- 🕒 3:23 ¶5, Easy to use too in Six Pack in 30 Days
- 🕒 3:26 ¶8, Easy to use in Six Pack in 30 Days
- 🕒 5:10 ¶5, Super easy and fun to use in Headspace
- 🕒 5:21 ¶9, App is smooth to use in Headspace
- 🕒 6:11 ¶8, it's extremely convenient and easy to use. in YAZIO Calorie Counter & Intermittent Fasting App
- 🕒 6:15 ¶10, this is much easier and intuitive in YAZIO Calorie Counter & Intermittent Fasting App
- 🕒 7:1 ¶1, It's so simple & easy to operate in Step Tracker - Pedometer
- 🕒 7:7 ¶4, It is very easy to use in Step Tracker - Pedometer
- 🕒 7:11 ¶6, I like the ease of the app to keep you updated with your current walking in Step Tracker - Pedometer
- 🕒 7:13 ¶7, Easy way to track your steps in Step Tracker - Pedometer
- 🕒 7:28 ¶15, I love that i don't have to remember to turn it on before hand and it continuously keeps track for m.... in Step Tracker - Pedometer

Groups:

- 👤 Good design (related to UI and UX)
-

- **Good design (related to UI and UX): Good for beginners**

Quotations:

- 🕒 5:2 ¶1, The meditation exercises are great for beginners like me in Headspace
- 🕒 5:15 ¶6, I love the fact that you have so much structure for a beginner. in Headspace

Groups:

- 👤 Good design (related to UI and UX)
-

- **Good design (related to UI and UX): links to additional material**

Quotations:

- 🕒 3:34 ¶14, There are links to videos too in case you want to see a real person demonstrate in Six Pack in 30 Days

Groups:

- 👤 Good design (related to UI and UX)
-

- **Good design (related to UI and UX) : does not need manual start or set up**

Quotations:

- 🕒 7:28 ¶15, I love that i don't have to remember to turn it on before hand and it continuously keeps track for m... in Step Tracker - Pedometer
- 🕒 7:38 ¶20, without setting anything up in Step Tracker - Pedometer

Groups:

- 👤 Good design (related to UI and UX)
-

- **Good design (related to UI and UX): additional feaures: easy to read charts**

Quotations:

- 🕒 7:4 ¶1, has easy-to-read charts to help you visualize your progress. in Step Tracker - Pedometer

Groups:

- 👤 Good design (related to UI and UX)
-

- **Good design (related to UI and UX): Easy to undertand**

Quotations:

- 🕒 3:17 ¶2, understand in Six Pack in 30 Days
- 🕒 3:33 ¶14, the exercises are easy to learn and are demonstrated pretty well in Six Pack in 30 Days
- 🕒 5:47 ¶6, The design is intuitive and easy to use in Headspace
- 🕒 7:8 ¶4, The layout is clear and concise in Step Tracker - Pedometer
- 🕒 7:12 ¶1, easy-to-read charts to help you visualize in Step Tracker - Pedometer

Groups:

- 👤 Good design (related to UI and UX)
-

- **Good design (related to UI and UX): multiple functionalities**

Quotations:

- 🕒 1:39 ¶1, I love the Google Fit app for its many different kinds of workout trackers in Google Fit
- 🕒 1:48 ¶5, it's the only app you can track these range of workouts with in Google Fit
- 🕒 7:9 ¶4, gives you steps and calories burned in Step Tracker - Pedometer
- 🕒 7:26 ¶14, Not only does it calculate steps and miles, it also gives attainable goals to reach in Step Tracker - Pedometer
- 🕒 7:37 ¶20, Counts steps, time, distance, and calories burned in Step Tracker - Pedometer
- 🕒 8:17 ¶8, -Very detailed, professional, most importantly, information is the very latest available in WebMD
- 🕒 8:35 ¶17, It's versatile to many ailments and conditions in WebMD

Groups:

- 👤 Good design (related to UI and UX)
-

- **Methods the user used to overcome the problems: adding permissions**

Quotations:

- 🕒 1:73 ¶18, allowed every permission possible in Google Fit

Groups:

- 👤 Methods the user used to overcome the problems
-

- **Methods the user used to overcome the problems: changing the password**

Quotations:

- 🕒 5:30 ¶12, Try changing the password linked to my email in Headspace

Groups:

- 👤 Methods the user used to overcome the problems
-

- **Methods the user used to overcome the problems: clearing the cache**

Quotations:

- 🕒 1:7 ¶3, I've logged out, back in, cleared cache, restarted phone and watch, uninstalled, reinstalled, and ev... in Google Fit
- 🕒 1:41 ¶4, I've cleared the cache in Google Fit
- 🕒 1:53 ¶7, clearing the cache in Google Fit
- 🕒 1:69 ¶18, Cleared cache in Google Fit

Groups:

- 👤 Methods the user used to overcome the problems
-

• **Methods the user used to overcome the problems: clearing the storage or the data**

Quotations:

- 1:43 ¶ 4, cleared storage in Google Fit
- 1:74 ¶ 18, Cleared data in Google Fit

Groups:

- Methods the user used to overcome the problems
-

• **Methods the user used to overcome the problems: close and relaunch**

Quotations:

- 5:26 ¶ 10, I have to close out and relaunch in Headspace

Groups:

- Methods the user used to overcome the problems
-

• **Methods the user used to overcome the problems: logging out and the logging back in**

Quotations:

- 1:7 ¶ 3, I've logged out, back in, cleared cache, restarted phone and watch, uninstalled, reinstalled, and ev... in Google Fit

Groups:

- Methods the user used to overcome the problems
-

• **Methods the user used to overcome the problems: rebooting**

Quotations:

- 7:32 ¶ 17, I uninstalled, reinstalled Step Tracker and rebooted in Step Tracker - Pedometer

Groups:

- Methods the user used to overcome the problems
-

• **Missing feature (Failing to Communicate with other devices and apps): no follow up apps**

Quotations:

- 3:31 ¶ 12, I do wish there was an app afterwards to follow with ab training because in Six Pack in 30 Days

Groups:

- Missing feature (Failing to Communicate with other devices and apps)
-

• **Missing feature (Failing to Communicate with other devices and apps): no google cast**

Quotations:

- 5:43 ¶ 18, No Google casting, in Headspace

Groups:

- Missing feature (Failing to Communicate with other devices and apps)
-

• **Missing feature (Failing to Communicate with other devices and apps): not communicating with a third party step counter app**

Quotations:

- 1:31 ¶ 16, I only use this to pair with an app that uses the step counter in Google Fit

Groups:

- Missing feature (Failing to Communicate with other devices and apps)
-

• **Missing feature (Failing to Communicate with other devices and apps): not communicating with fossil watch**

Quotations:

- 1:65 ¶ 13, The data from my fossil watch never syncs in Google Fit

Groups:

- Missing feature (Failing to Communicate with other devices and apps)
-

• **Methods the user used to overcome the problems: restarting the device**

Quotations:

- 1:2 ¶ 1, fix it by restarting the watch in Google Fit
- 1:7 ¶ 3, I've logged out, back in, cleared cache, restarted phone and watch, uninstalled, reinstalled, and ev... in Google Fit
- 1:54 ¶ 7, restarting my phone and watch in Google Fit

Groups:

- Methods the user used to overcome the problems
-

• **Methods the user used to overcome the problems: turning off battery optimization**

Quotations:

- 1:72 ¶ 18, turned off battery optimisation in Google Fit

Groups:

- Methods the user used to overcome the problems
-

• **Methods the user used to overcome the problems: uninstall and reinstall**

Quotations:

- 1:7 ¶ 3, I've logged out, back in, cleared cache, restarted phone and watch, uninstalled, reinstalled, and ev... in Google Fit
- 1:42 ¶ 4, uninstalled and reinstalled the app in Google Fit
- 1:52 ¶ 7, I have already tried uninstalling and reinstalling the app in Google Fit
- 1:67 ¶ 13, I even reinstalled to see if the data sync in Google Fit
- 1:70 ¶ 18, uninstalled, reinstalled, then uninstalled reinstalled in Google Fit
- 5:37 ¶ 15, I have Uninstalled and reinstalled once in Headspace
- 7:31 ¶ 17, I uninstalled, reinstalled Step Tracker in Step Tracker - Pedometer

Groups:

- Methods the user used to overcome the problems
-

• **Methods the user used to overcome the problems: using GPS training**

Quotations:

- 7:24 ¶ 12, I had to use the GPS training in Step Tracker - Pedometer

Groups:

- Methods the user used to overcome the problems
-

• **Missing feature (Failing to Communicate with other devices and apps): not communicating with samsung galaxy watch 3**

Quotations:

- 1:78 ¶ 10, Can't put Fit on my watch either in Google Fit

Groups:

- Missing feature (Failing to Communicate with other devices and apps)
-

• **Missing feature (Failing to Communicate with other devices and apps): not communicating with samsung watch4**

Quotations:

- 1:1 ¶ 1, on my Samsung Watch 4 with Google Fit and then Samsung altogether stops tracking my steps and calories in Google Fit

Groups:

- Missing feature (Failing to Communicate with other devices and apps)
-

• **Missing feature (Failing to Communicate with other devices and apps): not communicating with Suunto watch7**

Quotations:

- 1:15 ¶ 7, stops working on my Suunto 7 watch, and on my phone. in Google Fit

Groups:

- Missing feature (Failing to Communicate with other devices and apps)
-

• **Missing feature (Failing to Communicate with other devices and apps): not compatible with Fitbit**

Quotations:

- 6:28 ¶ 17, but is not compatible with Fitbit in YAZIO Calorie Counter & Intermittent Fasting App

Groups:

- Missing feature (Failing to Communicate with other devices and apps)
-

● **Missing feature (Failling to Communicate with other devices and apps): not working properly with Walgreens app**

Quotations:

🗨 1:79 ¶10, Walgreen can't be edited so no Rewards in Google Fit

Groups:

👉 Missing feature (Failling to Communicate with other devices and apps)

● **Missing feature (system functions): : being slow**

Quotations:

🗨 5:18 ¶7, It's also quite slow at remembering where I am in a program and general use in Headspace

Groups:

👉 Smart Component (lack of usefulness or accuracy) 👉 Smart component (usefulness and accuracy)

● **Missing feature (system functions): most options missing**

Quotations:

🗨 7:35 ¶18, A fundamental feature of the app (watch previous day's result) is no longer accessible, in Step Tracker - Pedometer

Groups:

👉 Methods the user used to overcome the problems 👉 Missing feature (system functions)

● **Missing feature (system functions): no cache**

Quotations:

🗨 6:8 ¶5, Could you please add a cache so that my frequent searches don't make a network request every time in YAZIO Calorie Counter & Intermittent Fasting App

Groups:

👉 Methods the user used to overcome the problems 👉 Missing feature (system functions)

● **Other disadvantages: too many ads**

Quotations:

🗨 8:1 ¶1, Home page looks like ad flyer in WebMD
 🗨 8:9 ¶4, The ad screen seems to be in the way in WebMD
 🗨 8:19 ¶9, Way too many ads in WebMD

Groups:

👉 Other advantages and disadvantages

● **Other disadvantages:Lack of easy cancelling options**

Quotations:

🗨 6:25 ¶16, You and I know that I may forget to cancel subscription once I stop using the app in a year. Why tak... in YAZIO Calorie Counter & Intermittent Fasting App

Groups:

👉 Bad design (lack of customization)

○ **Smart Component (health fuctions): Helps with remembering the medication**

Quotations:

🗨 8:18 ¶9, It helps me remember my meds in WebMD

Groups:

👉 Good design (related to smart component)

○ **Smart Component (health functions): adds saved meals quickly**

Quotations:

🗨 6:17 ¶10, adding a whole saved meal at once. in YAZIO Calorie Counter & Intermittent Fasting App

Groups:

👉 Good design (related to smart component)

● **Missing feature: failure to recieve feed back**

Quotations:

🗨 1:5 ¶2, Don't ask me to contact you in Google Fit
 🗨 1:8 ¶3, I've sent feedback through the app multiple times also in Google Fit
 🗨 1:11 ¶4, reported the issue as a bug and nothing has changed in Google Fit
 🗨 1:13 ¶6, The automatic response doesn't answer my point in Google Fit
 🗨 1:33 ¶17, There has been no explantation and no feedback as to a resolution in Google Fit
 🗨 1:71 ¶18, Sent feedback from inside the app time and time again. No feedback from Google in Google Fit

Groups:

👉 Methods the user used to overcome the problems 👉 Missing feature (system functions)

● **Other advantages: good free option**

Quotations:

🗨 3:24 ¶7, i am very suprising that it was free. in Six Pack in 30 Days
 🗨 6:29 ¶18, I didn't pay for pro but the free is just as good in YAZIO Calorie Counter & Intermittent Fasting App

Groups:

👉 Other advantages and disadvantages

● **Other disadvantages: Lack of free option**

Quotations:

🗨 5:3 ¶1, If we could access premium features by having ads and/or watching videos, that would be awesome! The... in Headspace
 🗨 5:5 ¶3, all functions are locked and there is a free trial period, but first give us your credit info. I'm s... in Headspace
 🗨 5:6 ¶4, nothings free anymore you need to upgrade to be able to use any of the app in Headspace
 🗨 5:9 ¶4, but now I need to upgrade to premium to get a decent night's sleep. in Headspace
 🗨 5:35 ¶14, I can't afford a membership and without it, even the basics are blocked.. in Headspace
 🗨 6:1 ¶1, You can't even start a fast without paying firs in YAZIO Calorie Counter & Intermittent Fasting App
 🗨 6:14 ¶9, The only way to modify any setting is to buy the pro version in YAZIO Calorie Counter & Intermittent Fasting App

Groups:

👉 Other advantages and disadvantages

○ **Smart Component (health functions): Helps being more focused**

Quotations:

🗨 5:11 ¶5, Highly recommend for stress, anxiety and focus! they also address pain, anger management, depression... in Headspace
 🗨 5:20 ¶9, Helps me concentrate when I need to in Headspace

Groups:

👉 Good design (related to smart component)

● **Smart Component (health functions): Helps to be more relaxed**

Quotations:

🗨 5:4 ¶1, I def do feel more relaxed physically & mentally/emotionally after doing a few exercises in Headspace
 🗨 5:11 ¶5, Highly recommend for stress, anxiety and focus! they also address pain, anger management, depression... in Headspace
 🗨 5:22 ¶9, allows you to keep track of your stress in Headspace

Groups:

👉 Good design (related to smart component)

● **Smart Component (health functions): Helps to fall asleep**

Quotations:

🗨 5:7 ¶4, have helped me to get to sleep in Headspace
 🗨 5:8 ¶2, Last night I put on a story called "Timeless Clockshop". I don't know how long I had been listening... in Headspace
 🗨 5:16 ¶6, The sleep casts are AMAZE-BALLS! I swear I'm not kidding - I haven't been able to stay awake until t... in Headspace
 🗨 5:27 ¶11, the Sleepcasts always work for me, I've never been awake at the end of any of them. in Headspace
 🗨 5:34 ¶14, could get me to sleep in such a way my nightmares were either non-existent, or just mild bad dreams.... in Headspace

Groups:

👉 Good design (related to smart component)

● Smart Component (health functions): Helps with anger management

Quotations:

- 5:11 ¶5, Highly recommend for stress, anxiety and focus! they also address pain, anger management, depression... in Headspace

Groups:

- Good design (related to smart component)
-

● Smart Component (health functions): Helps with depression

Quotations:

- 5:11 ¶5, Highly recommend for stress, anxiety and focus! they also address pain, anger management, depression... in Headspace

Groups:

- Good design (related to smart component)

● Smart Component (health functions): Helps with pain

Quotations:

- 5:11 ¶5, Highly recommend for stress, anxiety and focus! they also address pain, anger management, depression... in Headspace

Groups:

- Good design (related to smart component)

● Smart Component (health functions): Motivational

Quotations:

- 7:14 ¶9, My first day trying this app and it gave me encouragement to get off the couch in Step Tracker - Pedometer
- 7:21 ¶12, I love the achievements! Gives me incentive in Step Tracker - Pedometer
- 8:41 ¶19, I find interest in learning something new in WebMD

Groups:

- good design (other aspects)

- 1:27 ¶14, it is starting to mess up with distances walked and where I walked! I know the approximate distances... in Google Fit
- 1:30 ¶16, it will count my steps when I'm standing still but not when I'm actually walking. I just lost 90 min... in Google Fit
- 1:44 ¶14, Fit has stopped tracking ANY steps in Google Fit
- 1:55 ¶8, Very often the app will not read data from the phone properly and may steel up to 5000 steps in a da... in Google Fit
- 1:61 ¶11, It is not tracking steps, km , heart rate etc in Google Fit
- 1:64 ¶12, it is wrong 100 percent of the time. It usually gets about half sometimes less of the steps I actual... in Google Fit
- 1:68 ¶15, Unreliable and inaccurate tracking. Heart points disappear all of a sudden. On the day of workout it... in Google Fit
- 1:76 ¶19, The paced walking isn't tracking correctly. If I set it to 100 steps per minute i expect it to read... in Google Fit
- 5:33 ¶13, the issue was that when selecting the time, a different one played, so a 10min one would last 16min... in Headspace
- 6:3 ¶2, I also have to edit foods frequently so that their nutrition facts match those listed on the packagi... in YAZIO Calorie Counter & Intermittent Fasting App
- 6:23 ¶14, app appears to be counting calories well and not any of the vitamins or minerals in YAZIO Calorie Counter & Intermittent Fasting App
- 7:5 ¶2, -App adds thousands of steps i haven't walked. Very frustrating for anyone genuinely trying to track... in Step Tracker - Pedometer
- 7:16 ¶10, I take the same route every day, and one day it shows I walked 5.1 miles, and the next it can show 6... in Step Tracker - Pedometer
- 7:17 ¶10, And the map that shows the route I walked is way off too in Step Tracker - Pedometer
- 7:19 ¶11, When I use the tracker to track my walk, it adds way too many steps and miles I didn't actually comp... in Step Tracker - Pedometer
- 7:22 ¶12, Lately it has been losing steps in Step Tracker - Pedometer
- 7:23 ¶12, Now it is losing distance in Step Tracker - Pedometer
- 7:25 ¶13, Sometimes fails to accurately track mileage in Step Tracker - Pedometer
- 7:30 ¶17, GPS TRACKING IS STILL BROKEN in Step Tracker - Pedometer
- 7:34 ¶17, GPS and mileage are way off in Step Tracker - Pedometer
- 8:12 ¶5, This app is terrible for people who have multiple symptoms or diagnosis. It doesn't save any of your... in WebMD
- 8:21 ¶11, Having pain in abdomen, and lower back, can't stand or sit for more than 2 minutes, was told maybe k... in WebMD
- 8:27 ¶14, This new form says (Over Exaggeration) that I died 7 years ago. in WebMD

Groups:

- Smart Component (lack of usefulness or accuracy)

○ Smart Component (high level of customization): Being able to choose the length of the session

Quotations:

- 5:1 ¶1, being able to choose how long they last is great in Headspace

Groups:

- Good design(high level of customization)

○ Smart Component (high level of customization):being able to choose the difficulty level

Quotations:

- 3:11 ¶9, Perfect app for all levels of fitness in Six Pack in 30 Days
- 3:14 ¶1, you can adjust the exercises difficulty which is very helpful in Six Pack in 30 Days
- 3:21 ¶14, want more hard then u can choose it in Six Pack in 30 Days
- 3:37 ¶18, Good it asks you when to remind you it makes it easier if the exercise is to hard and give in Six Pack in 30 Days

Groups:

- Good design(high level of customization)

● Smart Component (high level of customization):feedback based plans

Quotations:

- 3:19 ¶3, The coach system can adjust the workouts based in the feedback you give in Six Pack in 30 Days

Groups:

- Good design(high level of customization)

● Smart Component (lack of usefulness or accuracy): miscalculation

Quotations:

- 1:3 ¶2, my run distances go down from 5km (the exact distance I always do. It's correct) to between 2.5km an... in Google Fit
- 1:6 ¶3, Heart points consistently disappear through out the day. Sometimes they come back in Google Fit
- 1:12 ¶6, 5-8 hour bike trips give 1 heart point, while a walk for 30 to 50 minutes gives 13-15 points. in Google Fit
- 1:16 ¶7, my heart rate is not accurate and the amount of calories burned is way off. in Google Fit
- 1:20 ¶8, to work properly, I have to close it every time as it seems to not read data otherwise. The phone se... in Google Fit

● Smart Component (scientific accuracy): being scientifically reliable

Quotations:

- 3:38 ¶19, systemetic and scientific way of reducing your fat and get a tuned body in Six Pack in 30 Days

Groups:

- good design (other aspects)

● Smart Component (scientific accuracy): informative

Quotations:

- 3:35 ¶15, Great App Saves time by letting you know what exercises and time required in Six Pack in 30 Days
- 8:17 ¶8, -Very detailed, professional, most importantly, information is the very latest available in WebMD
- 8:36 ¶17, Saves a trip to the Dr or at least gives you symptons you can arm yourself with before going in WebMD
- 8:37 ¶18, Great health info in WebMD
- 8:40 ¶19, Great information so helpful in WebMD

Groups:

- good design (other aspects)

○ Smart Component (usefulness and accuracy):Being able to search specific programs

Quotations:

- 5:24 ¶9, you could search something more specific in Headspace

Groups:

- Good design(high level of customization)

● Smart component(usefulness and accuracy): feels like a real coach

Quotations:

- 3:36 ¶16, You have your virtual trainer with your as if you are in a gymnasium in Six Pack in 30 Days

Groups:

- Smart component (usefulness and accuracy)

• **Smart component(usefulness and accuracy):adjust plans based on the feedback**

Quotations:

- 5:19 ¶3, The coach system can adjust the workouts based in the feedback you give in Six Pack in 30 Days

Groups:

- Smart component (usefulness and accuracy)
-

• **Smart component(usefulness and accuracy):breaking down portions into ingredients**

Quotations:

- 6:30 ¶18, it even breaks down if you just have portions. in YAZIO Calorie Counter & Intermittent Fasting App

Groups:

- Smart component (usefulness and accuracy)
-

• **Smart component(usefulness and accuracy):good measurments**

Quotations:

- 7:3 ¶11, It measures and records everything from time to distance and steps in Step Tracker - Pedometer
- 7:9 ¶14, gives you steps and calories burned in Step Tracker - Pedometer
- 7:15 ¶18, Great app for tracking all steps, even ones you don't realize you're taking in Step Tracker - Pedometer
- 7:27 ¶14, this app couples the distance that one achieves with mile markers such as 26 which equals the mile d... in Step Tracker - Pedometer
- 7:33 ¶17, Step counts, time recording, and calories burned are accurate in Step Tracker - Pedometer
- 7:36 ¶19, I added two step tracking apps and they both stay within a few steps of each other so very happy thi... in Step Tracker - Pedometer
- 7:39 ¶20, And it works! Like, really well. in Step Tracker - Pedometer

Groups:

- Smart component (usefulness and accuracy)
-

◦ **Smart Component(high level of customization):varied options**

Quotations:

- 3:15 ¶11, it gives you a variety of different good exercises that should take you about 20 minutes to do in Six Pack in 30 Days
- 3:25 ¶17, it is in three modes for belly fat loses and six pack gainers and who want to do the packs to rock a... in Six Pack in 30 Days
- 3:29 ¶10, This app is not for just belly, its also full body workout in Six Pack in 30 Days
- 3:30 ¶111, I don't use it to get a six pack but to get a good core work out in Six Pack in 30 Days
- 3:39 ¶20, choose what you want and in general in Six Pack in 30 Days
- 5:19 ¶18, This app will do wonders for you if you let it. Make it a part of your daily routine and you will be... in Headspace
- 5:31 ¶13, Generally a good and varied selection of meditations and of times. in Headspace
- 5:38 ¶16, They have a little bit of everything, from beginners courses to more advanced, meditations to do whi... in Headspace
- 6:27 ¶17, amount of products in YAZIO Calorie Counter & Intermittent Fasting App
- 6:34 ¶20, I can track my fasting, calories, and activities all in one app! in YAZIO Calorie Counter & Intermittent Fasting App
- 7:29 ¶16, I really appreciate being able to change the sensitivity level in the app in Step Tracker - Pedometer

Groups:

- Good design(high level of customization)
-

◦ **Smart Component(lack of customization): does not track carbs**

Quotations:

- 6:10 ¶7, Useless for Keto, even with a Pro subscription. It will track Carbs & dietary fiber, but not sugar a... in YAZIO Calorie Counter & Intermittent Fasting App

Groups:

- Methods the user used to overcome the problems
 - Missing feature (system functions)
-

◦ **Smart Component(lack of customization): lack of serving options**

Quotations:

- 6:16 ¶10, that there should be more options for a standard serving when adding new gigs than ounces and fluid... in YAZIO Calorie Counter & Intermittent Fasting App
- 6:19 ¶12, Serving sizes need to be Imperial and Metric in YAZIO Calorie Counter & Intermittent Fasting App

Groups:

- Bad design (lack of customization)
-

• **Smart component(usefulness and accuracy):good suggestions**

Quotations:

- 5:23 ¶9, then suggests programs that could help at the moment. Has different things to do depending of the ti... in Headspace

Groups:

- Smart component (usefulness and accuracy)
-

• **Smart component(usefulness and accuracy):track steps**

Quotations:

- 5:39 ¶16, these guided meditations in Headspace
- 7:6 ¶3, I can keep track of my steps and miles and it come in handy during my workouts in Step Tracker - Pedometer

Groups:

- Smart component (usefulness and accuracy)
-

• **Smart component(usefulness and accuracy):tracks the workout**

Quotations:

- 3:27 ¶18, track your work out in Six Pack in 30 Days

Groups:

- Smart component (usefulness and accuracy)
-

◦ **Smart Component(lack of customization): no vegan plan**

Quotations:

- 6:12 ¶9, There should be an option when setting things up to select what kind of diet you eat. For someone wh... in YAZIO Calorie Counter & Intermittent Fasting App

Groups:

- Bad design (lack of customization)
-

◦ **Smart Component(lack of customization): no way to set the desired weigh loss and the pace**

Quotations:

- 6:13 ¶9, Also I didn't see an option to change how much you wanted to lose and how fast in YAZIO Calorie Counter & Intermittent Fasting App

Groups:

- Bad design (lack of customization)
-

◦ **Smart Component(lack of customization):does not allow editing**

Quotations:

- 6:4 ¶3, doesn't allow editing of scannable ingredients in YAZIO Calorie Counter & Intermittent Fasting App
- 8:3 ¶1, It let me search 2 meds to add to her profile. Once set, I found no way to add, edit, or remove a sc... in WebMD

Groups:

- Bad design (lack of customization)
-

• **Smart component(lack of usefulness and accuracy : Failing To Show Information- not saving information**

Quotations:

- 1:46 ¶5, saving never works in Google Fit
- 1:49 ¶18, Lately most activities aren't recorded at all anymore in Google Fit
- 1:51 ¶7, It literally won't track any activities in Google Fit

Groups:

- Wrong information
-

• **Smart component(lack of usefulness and accuracy): deleting both posts when one is deleted**

Quotations:

- 8:16 ¶7, it posts two posts so of course, I delete one which apparently delete both in WebMD

Groups:

- Wrong information

• **Smart component(lack of usefulness and accuracy): failing to load the paid material even when subscribed**

Quotations:

- 5:29 ¶12, I have a subscription, and yet, everything thing is locked away. When I go to subscribe again, it te... in Headspace

Groups:

- Wrong information

• **Smart component(lack of usefulness and accuracy): Failing To Show Information**

Quotations:

- 1:25 ¶11, everything is showing zero. in Google Fit
- 1:29 ¶15, Heart points disappear all of a sudden in Google Fit
- 1:32 ¶17, this app stopped scoring "Heart Points" for workout activities that are manually entered such as an... in Google Fit
- 1:47 ¶5, Often it's need a long time to load, but at the nothing is happening in Google Fit
- 1:62 ¶11, I took a walk for more than hr but application never captured any data in Google Fit
- 8:20 ¶10, Did a simple search for the OTC vitamin Niacin and every single category just says "consult your pha... in WebMD

Groups:

- Wrong information

• **Smart Coponent (Lack of usefulness and accuracy): limited data base**

Quotations:

- 6:9 ¶6, you can't input all nutrients, such as potassium in YAZIO Calorie Counter & Intermittent Fasting App
- 6:24 ¶15, The bar code scanner doesn't recognize most of the common foods as well in YAZIO Calorie Counter & Intermittent Fasting App
- 6:31 ¶19, There is a shortage of a lot of foods in YAZIO Calorie Counter & Intermittent Fasting App

Groups:

- Methods the user used to overcome the problems
- Missing feature (system functions)

• **Smart Coponent(lack of customization):lack of options**

Quotations:

- 1:58 ¶9, heart points seem to be a "priority" in the app when I prefer to always look at my steps in Google Fit
- 6:33 ¶19, and it lacks some of the features in the subscription. in YAZIO Calorie Counter & Intermittent Fasting App
- 8:3 ¶1, It let me search 2 meds to add to her profile. Once set, I found no way to add, edit, or remove a sc... in WebMD
- 8:5 ¶3, YOU TOOK AWAY NEARLY ALL OF THE CUSTOM FEATURES in WebMD
- 8:6 ¶3, I CAN'T EVEN PICK AT WHAT LEVEL TO BE NOTIFIED - THIS ONE ONLY ALLOWS NOTICES WHEN LEVELS ARE HIGH.... in WebMD
- 8:24 ¶13, This app only allows you to select pre-set brand name pharmaceuticals. Cannot add your own custom it... in WebMD
- 8:39 ¶18, I wish there was a way to select all at once. Maybe create a group of medications in WebMD

Groups:

- Bad design (lack of customization)

• **system issues (Failure of unknown origin) :Failing To Show Information - failing to load indormation on a new phone**

Quotations:

- 5:41 ¶17, got a new phone and cannot log into my account anymore, in Headspace

Groups:

- Wrong information

• **Smart component(lack of usefulness and accuracy): Failing To Show Information- no past data visible**

Quotations:

- 1:9 ¶4, I can no longer access past data in Google Fit

Groups:

- Wrong information

• **Smart component(lack of usefulness and accuracy): not offering many choices when uncertain**

Quotations:

- 8:13 ¶5, If it does save any, it doesn't give you any possible results in WebMD

Groups:

- Smart Component (lack of usefulness or accuracy)

• **Smart component(lack of usefulness and accuracy): showing wrong information)**

Quotations:

- 8:42 ¶20, WebMD is sharing information which is misleading, but mostly completely false in WebMD

Groups:

- Wrong information

• **Smart component(lack of usefulness and accuracy: Failing To Show Information- Failling to load the downloaded content offline**

Quotations:

- 5:17 ¶7, The application will not store or remember that I have downloaded a program to be able to use offlin... in Headspace

Groups:

- Wrong information

• **system issues (Failure of unknown origin)): Failing To Show Information-failing to load anything**

Quotations:

- 5:45 ¶20, Changed to 1 star because nothing will load in Headspace

Groups:

- Wrong information

• **system issues (Failure of unknown origin): app stops or does not start to work**

Quotations:

- 1:75 ¶18, Doesn't work in Google Fit
- 5:25 ¶10, but recently there's an issue on the app where sleep music will stop, usually right around 5 minutes... in Headspace
- 5:28 ¶11, BUT there isn't much more frustrating then listening to a sleepcast, just dozing off, to find out ha... in Headspace
- 5:32 ¶13, it so often stops working for me in Headspace
- 5:44 ¶19, Every meditation, sleep cast or music that I listen to constantly stops / drops out and then (mostly... in Headspace
- 5:46 ¶20, The app no longer seamlessly streams the meditations in Headspace

Groups:

- System issues

• **System issues (Failure of unknown origin): error while working on phone**

Quotations:

- 5:36 ¶15, app has an error every few minutes during mediation on my phone in Headspace

Groups:

- System issues

• **System issues (Failure of unknown origin): reset does not work**

Quotations:

- 6:18 ¶11, Reset Doesn't Work in YAZIO Calorie Counter & Intermittent Fasting App

Groups:

- System issues

• **System issues (Failure of unknown origin): today tab disappears**

Quotations:

🕒 5:42 ¶ 18, The "today" tab often just disappears in Headspace

Groups:

👁 System issues

• **System issues (other aspects):loads slowly**

Quotations:

🕒 1:45 ¶ 5, it's need a long time to load in Google Fit
🕒 6:7 ¶ 5, feels slow in YAZIO Calorie Counter & Intermittent Fasting App
🕒 6:21 ¶ 13, loads very very slowly. in YAZIO Calorie Counter & Intermittent Fasting App

Groups:

👁 System issues

◦ **TITLE**

Quotations:

🕒 1:80 ¶ 3, When it works it's great. in Google Fit
🕒 1:81 ¶ 5, Oops, something went rong in Google Fit
🕒 5:48 ¶ 7, Content is great, but the app is poor. in Headspace
🕒 5:49 ¶ 12, Fix your bugs in Headspace
🕒 5:50 ¶ 18, Content is great, but the app needs some work in Headspace
🕒 7:40 ¶ 2, thousands of steps i haven't walked in Step Tracker - Pedometer
🕒 7:41 ¶ 3, This app literally does it all, in Step Tracker - Pedometer
🕒 7:42 ¶ 17, I don't understand what happened in Step Tracker - Pedometer
🕒 7:43 ¶ 20, Why can't all apps just work like this? in Step Tracker - Pedometer