universität wien

MASTERARBEIT / MASTER'S THESIS

Titel der Masterarbeit / Title of the Master's Thesis

„A Conversational Agent for Positive Change
A personalityzation approach“

verfasst von / submitted by

Oliver Guggenberger

angestrebter akademischer Grad / in partial fulfilment of the requirements for the degree of

Master of Science (MSc)

Wien, 2020 / Vienna, 2020

Studienkennzahl lt. Studienblatt / degree programme code as it appears on the student record sheet:	UA 066 935
Studienrichtung lt. Studienblatt / degree programme as it appears on the student record sheet:	Master Medieninformatik
Betreut von / Supervisor:	Univ.-Prof. Dipl.-Math. Dr. Peter Reichl

Acknowledgments

I would like to express my highest appreciation to the following persons who greatly supported me during this thesis project:

My supervisor Univ.-Prof. Dipl.-Math. Dr. Peter Reichl, for making this project possible. His door was always open and he was always very accommodating. Thank you!

My co-supervisor Dipl-Ing. Christian Löw, for his patient guidance and valuable advice throughout our many meetings. Thank you for answering all my questions and for your extensive support.

I would also like to thank the participants of the user study for your time and valuable feedback. Thank you for staying on and supporting my research. Without your passionate participation and input, the user study could not have been successfully conducted.

Finally, I want to express my very profound gratitude to my parents for providing me with your unfailing support and continuous encouragement throughout my years of study and through the process of researching and writing this thesis. This accomplishment would not have been possible without you. Thank you!

Oliver Guggenberger

Abstract

This study investigates the effects of applying the Positive Computing paradigm to conversational agents by utilizing a personalityzation approach. Although the importance of the user's well-being is recognized by researchers, the application of the Positive Computing framework to conversational agents has just begun. Related research has revealed that personalization techniques can improve user experience and overall user satisfaction. Facilitating a satisfying user experience is essential, especially in the context of motivational and behavioral change technologies. This paper examines how personalization and the application of Positive Computing can be brought together. Therefore, this thesis investigates how a personalityzation approach can facilitate the Positive Computing paradigm to increase the perceived competence to make or maintain a change towards a healthy behavior. Furthermore, the research interest is to find out whether the METUX model can be applied to conversational agents.

Findings from a previously conducted research survey were used to form the concept of this work. An existing health care chatbot was analyzed and revised. As a result, a disposition sensible health care chatbot was developed to test and evaluate the concept. A conceptually validating approach was applied based on a mixed-method approach. Thus, both quantitative and qualitative research methods were utilized in a participative user study. A personality test was conducted to acquire the personality of the respective users and to better understand their specific needs. The METUX model was used to structure and evaluate this project. Finally, semi-structured interviews were conducted. The evaluation of the user study indicates that participants perceived their competence to maintain a healthier lifestyle as increased after using the chatbot. Participants with an increased perceived competence score (PCS) rated their motivation to maintain a healthier lifestyle as rather autonomous compared to participants of a decreased PCS score. Conversely, participants perceived their competence to exercise regularly as being lower after having used the chatbot. Interestingly, these attendees rated their decision as rather autonomously taken. The analysis of the semi-structured interviews reveals that participants felt supported in their competence to achieve their health goals. They appreciated the features of the conversational agent and stated that their consciousness of healthy living had increased.

The findings indicate ambiguous results for the first research question. Thus, further research will be required with the help of a large-scale user study. Nevertheless, the adoption of the PC framework worked well in this study. The personalityzation approach augmented the PC paradigm successfully.

Keywords: Interaction Design, Conversational Interface, Positive Computing, Personalityzation

Contents

List of Figures

Listings

Acronyms

AI	Artificial Intelligence
API	Application Program Interface
BPN	Basic Psychological Needs
CA	Conversational Agent
CUI	Conversational User Interface
CET	Cognitive Evaluation Theory
CI	Conversational Interface
DSD	Disposition Sensible Design
METUX	Motivation, Engagement & Thriving in User Experience
NLP	Natural Language Processing
OCEAN	Openness, Conscientiousness, Extraversion, Agreeableness, Neuroticism
PC	Positive Computing
PCS	Perceived Competence Scale
PNSES	Psychological Need Satisfaction in Exercise Scale
TENS	Technology based Experience of Need Satisfaction
TDEE	Total Daily Energy Expenditure
TSRQ	Treatment Self-Regulation Questionnaire
UCD	User Centered Design
UX	User Experience
UI	User Interface
VUI	Voice User Interface
SDT	Self Determination Theory
SUS	System Usability Scale
SWB	Subjective Well-being

1 Introduction

This master thesis is the final part of the master of the media informatics program at the University of Vienna. Additionally, it is part of a PhD project pursued at the research group Cooperative Systems of the University of Vienna [39]. The work presented in this thesis was conducted and written between February 2019 and March 2020. The workload corresponds to 30 ECTS. In order to improve readability, only the male form is used in the text, nevertheless all data apply to members of both genders.

The first conversational agent ELIZA was created in 1966. In the following years, it became quiet around this topic. However, in spring 2016 conversational agents experienced a revival. Brandtzaeg and Følstad justify this development through massive advances in artificial intelligence (AI) and the extensive usage of messaging platforms like Facebook Messenger, Whatsapp and Telegram [8]. Today, at the beginning of the year 2020, chatbots are still a hot topic, although various challenges occurred and persist. Service providers and big companies like the big four (Microsoft, Amazon, Google and Facebook) see conversational interfaces as a new potential route to customer engagement [13], not only because of the huge user base of the respective messaging platforms, but also due to the ”savings” this technology can make for their companies [38]. Dale R. [13] stated that there are a total of 2.1 billion messaging users and one billion of those using the Facebook messenger. These impressive numbers and the general development in this field clearly indicate that conversational interfaces are still an emerging field. Despite this progress, the user’s perspective is still often neglected. Users often judge chatbots to be unintelligent and non-conversational.

In her master thesis Smestad [38] voiced that one thing about conversational agents we can improve is how they are perceived. As a result of her thesis, she developed a personality framework whose intention is to provide a stable pattern in how chatbots are perceived.

The idea to improve the user experience by adapting a piece of software to them is not new and has proved to be an effective way to increase the user acceptance and satisfaction [4], [35], [38]. Despite the awareness of the positive effects of personalization, few studies exist which investigate the effects of disposition on user interfaces and interaction design [35].

While personalization often refers to the customization of the user interface (UI) and the shallow setting of user preferences, the more specialized term personalityzation refers to the adaption of a computer system to the user personality. The term personalityzation was first introduced by Arazy [4] in order to distinguish from the broader term of personalization.

Interactive computer programs that are able to adapt to their users can give valuable support in achieving specified goals. Some goals are hard to achieve, especially health care goals which aim to facilitate behavioral change. Hence, those programs are designed to ease the challenges of

maintaining beneficial goals.

In psychology, various theories exist which explain the reason and motivation of human behavior (Self-Determination Theory, Theory of Planned Behavior, Attribution Theory, Nudge Theory, etc.). Those theories can be used to better understand the users and maintain and foster their motivation to use software.

The Positive Computing paradigm has its origin in the positive psychology movement and is concerned with the question of how to develop software which has a positive influence on its users or even on the entire world.

In this thesis, an attempt was made to leverage the Positive Computing framework in the context of conversational agents with respect to personalityzation. As a result of this considerations, a disposition sensible health care bot was developed, tested and subsequently evaluated.

1.1 Contribution

Through this thesis, a step was taken in the direction of chatbot personalization. An investigation of a disposition sensible design (DSD) for a conversational agent was made and contributes to the understanding of how personalized conversational agents can be developed. The Positive Computing paradigm was implemented to a DSD health care chatbot and subsequently evaluated through a participative user study. Additionally, a contribution was made to how a conversational agent, following the PC paradigm, can be evaluated. Furthermore, a framework obeying the Positive Computing paradigm was created for modelling and developing personalized conversational interfaces. The main contributions are summarized as follows:

1. Application of the Positive Computing paradigm to a disposition sensitive health care chatbot
2. Elaboration of a Positive Computing personalityzation framework
3. Development of a disposition sensitive health care chatbot
4. Evaluation of a disposition sensitive health care chatbot

1.2 Research Interests

Based on the insights of the research survey, following research questions were formulated:

RQ1 Can a chatbot augment/increase the perceived competence of its user in being able to make or maintain a change towards a healthy behavior?

(a) To what extent is a chatbot able to increase the perceived competence to make or maintain a change towards a healthy behavior?

RQ2 Can the Positive Computing framework/paradigm (METUX) be applied to conversational agents?

(a) Does disposition sensible design work in line or rather augment the Positive Computing paradigm?

1.3 Outline

This work consists of seven chapters and is structured as follows:

Chapter 1 introduces the topic and briefly describes conversational agents, Positive Computing, personalization and personalityzation. Furthermore, the link to the field of psychology is explained.

Chapter 2 describes the frame and the background of this thesis as well as the insights of the previously conducted survey paper. Additionally, challenges and risks are discussed.

Chapter 3 introduces the methodology used for this thesis and summarizes the research questions.

Chapter 4 presents the developed concept based on literature research and experiences of related research. It is about theories from the field of psychology and the way they are utilized for the concept of this work. Furthermore, it is about personalization and personalityzation and how personalityzation can be used to increase user satisfaction.

Chapter 5 reveals details about the prototype implemented for this thesis. It describes the conversational agent as an instantiation of the concept introduced in chapter four. This chapter also contains a listing of features, third party services used, an architectural overview and used technologies to realize the present project.

Chapter 6 interprets and discusses the results and the findings of this paper.

Chapter 7 Summarizes the results and concludes this thesis. Additionally, a possible outline for future work is presented.

2 Background & Related Work

In this chapter, both related work and relevant background are presented to the reader to gain a comprehensive understanding of the topic. An introduction to motivational theories is provided as well as a link to chatbots. Next, the Positive Computing paradigm and relevant research are discussed. Lastly, notable evaluation methods are listed and reviewed.

2.1 Motivational Technologies & Behavioral Change

Maintaining a healthy behavior towards a beneficial and sustainable behavioral change can be one of the hardest challenges one can face. Various services, applications and theories in the area of health and well-being already exist which aim to ease these challenges.

In the Precious [10] project an attempt was made to cover existing theories from the field of psychology in a holistic way. Empirical validated principles and guidelines were developed and derived to improve the design of adaptive and motivational health care technologies. As a result, a framework was developed, capable of bridging the areas of motivational theories and behavioral science, as well as considering gamification aspects. Although a strong emphasis on individual factors like cultural differences of the individual users, demographic data and different kind of motivations was placed, the personality of the user was not taken into account. While the focus of this project is the improvement of the user's well-being through preventive interventions, the Positive Computing framework has not been considered by the authors. Instead, components of the PC paradigm like the Self-Determination-Theory were elaborated directly. Applying chatbots for achieving behavioral change was out of the scope of the Precious project. Though, especially chatbots hold great potential in helping their user since they can act as a personal assistant or companion. Hence, chatbots could be a perfect tool to facilitate healthy living and could even support positive change.

Self-Determination Theory

The Self-Determination Theory (SDT) is a very influential and empirically sound theory that links human motivation, personality and optimal functioning.

The SDT defines three basic psychological needs (BPN) that are universal, innate and intrinsic. Those BPNs are *autonomy*, *competence* and *relatedness*. The SDT postulates that the fulfilment of the three BPNs are essential for psychological health, well-being and flourishing. Peters et al. [47] stated that fostering *autonomy* would increase engagement, supporting

competence would increase motivation and facilitating *relatedness* would enhance well-being.

In this theory, a differentiation between intrinsic and extrinsic (controlled) types of motivation has been made. While intrinsic motivation comes from internal sources without the need for external regulation, controlled motivation depends on an external regulation like rewards or penalties. Intrinsic motivation is the most self-determined type of motivation and is defined as engaging in an activity for the pleasure and satisfaction derived from the activity itself [46]. The types of motivation can be placed on a continuum from controlled to autonomous motivation (see figure 1). An external controlled motivation that highly aligns with one's personal goals can be similar to internal motivation. A third type of motivation was stated by the SDT that is a motivation without an intention to act (i.e. amotivation)[16].

The Cognitive Evaluation Theory (CET) is a sub-theory of the SDT which investigates the effects of external consequences on internal motivation. Research has revealed that intrinsically motivated behavior can be undermined by extrinsic rewards. This phenomenon is referred as motivational "crowding out" [14].

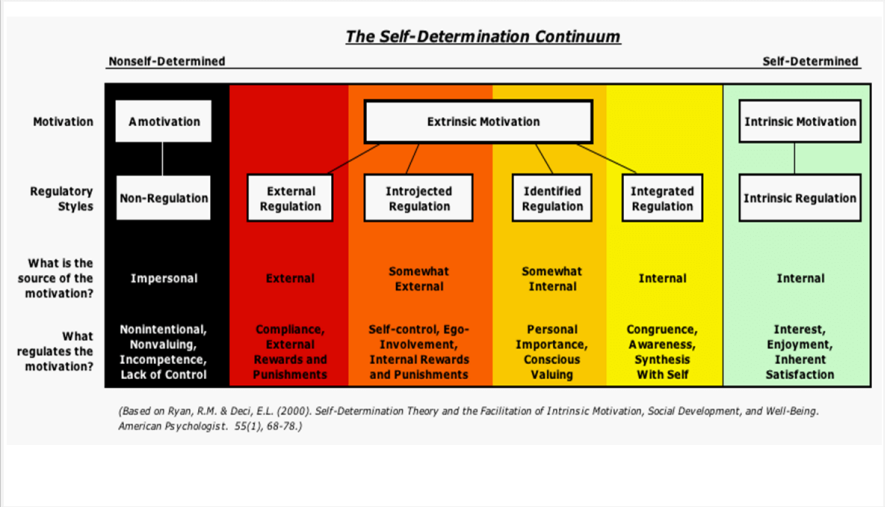


Figure 1: The Self-Determination Continuum[4]

Autonomy refers to the sense of feeling in control of yourself and to act in accordance with your own goals and interests. The theory states that

[4] `https://positivepsychology.com/self-determination-theory/`, *last accessed 2020-02-07*

an individual acting highly autonomously will show higher engagement and thus experience greater well-being.

The positive outcomes of *autonomy* and *competence* supportive design were evidenced in an interesting experiment by Peng et al. [46]. The authors investigated the effect of in-game features that foster *autonomy* and *competence* in an exergame. The participants of this experiment were presented different versions of the same game. In one experiment, the participants were asked to play two different versions of the same game. One version enabled *autonomy* supportive game features and one not. Subsequently, they were asked to rate their experience. The ratings clearly showed that the participants preferred the *autonomy* supportive version of the game with regard to game enjoyment, motivation for future play, likelihood of game recommendation and game rating.

Competence describes the need of feeling able and capable. It is a "need for challenge and feelings of effectance" [54]. Wilson et al. indicated a positive correlation between the need satisfaction of *competence* and more self-determined exercise regulations, which again correlates positively with exercise behavior, attitudes and physical fitness [63]. Peng et al. investigated the effect of *competence* supportive in-game features with the same method as described above. Participants preferred the version of the game with *competence* supportive features with regard to game enjoyment, motivation for future play, likelihood of game recommendation and game rating [46].

Relatedness refers to the psychological need to feel connected to others and is the core of most theories of well-being [6]. The psychological need for relatedness enhances well-being and intrinsic motivation when satisfied [46]. Ryan et al. [54] investigated in a study the extent to which Massive Multiplayer Online (MMO) games produce feelings of relatedness, and the motivational and well-being enhancements associated with it. Especially MMOs offers various opportunities for interaction between players and are suited to satisfy relatedness through a gaming session. The satisfaction of relatedness has shown to promote a sense of presence, game enjoyment and increased intention for future play.

The SDT is a broadly applied and well-accepted theory of human motivation which has been successfully applied in the context of computer programs. As SDT is also part of the PC paradigm, the application of the SDT theory through the PC paradigm on conversational agents is reasonable and promising to increase user satisfaction and well-being.

2.2 Chatbots

In 2016 we have been witnesses of the rebirth of chatbots. The first chatbot ELIZA appears in 1966. In between, user needs and motivations for using chatbots has changed [8]. In 2019 a total of 2.52 billion users accessed the top messaging platforms to communicate[5]. This indicates that people increasingly spend time on messaging platforms, which makes it very attractive for companies to invest in the development of chatbots. Despite the positive aspects of chatbots, there are some challenges that should be considered.

The use of chatbots does not merely imply a change of the user interface but also of the way users interact with services and data. The interaction has shifted from clicking buttons, swiping and scrolling on graphical user interfaces, to entering natural text into simple text fields or even no input on the interface (e.g. voice assistants like Google Assistant, Alexa or Siri). This is also referred as "Zero UI" where the interface is strongly or completely reduced. Therefore, the cognitive load a user has to manage is reduced, which in turn increases the user acceptance, willingness to interact with the technology and the overall user experience [19], [44].

Nonetheless, the user interface is for the bigger part a blank canvas where the contents and features of the offered service are basically hidden from the user. The vast variations of the user inputs and the openness of conversational interfaces are definitely the main challenges researchers and practitioners are facing. Valério et al. [62] analyzed different strategies to convey the features of a chatbot. These strategies should be considered in the service design of CAs to facilitate a better user experience.

One challenge for designers and researchers is, therefore, to see the conversation itself as the object of the design. However, the task to design technology is not anymore an explanatory task but rather an interpretational one [22]. In that sense, an explanatory task has the objective to explain to the user which features and content are available and which steps need to be taken to reach the desired goal through the interface. As opposed to this, an interpretational task has got the objective to understand the users, their desires, their goals and how the invisible features should be presented to them. Conversational breakdowns are identified by several authors as key challenge of chatbot development [22], [49]. Even with sophisticated chatbots like Google Assistant, the communication flow gets interrupted at some point. Facebook has recognised the problem of conversation breakdowns and argued that designers cannot expect perfection because people always behave in an unexpected manner [49]. Thus, especially AI-powered chatbots need substantial adaptation and maintenance to serve their task properly [8].

[5] `https://www.statista.com/statistics/483255/number-of-mobile-messaging-users-worldwide/`, *last accessed 2019-12-12*

Another major challenge for conversational agents is to keep and manage the context of the current conversation [49]. The research study of Cho J. [12] on a virtual home assistant (e.g. Google Home) has revealed that participants rather desire specific messages in case of losing the chatbot's context, than generic error messages like "I am not clear about your request". Jain et al. [33] have identified a discrepancy between the chatbot's state of understanding (context) and the users' perception of the chatbot's understanding. As a consequence, they proposed *Convey* (CONtextView), an additional graphical element, indicating the current chatbot's state. Participants preferred using chatbot's using *Convey* over chatbots without *Convey*.

Despite the importance of keeping the conversational flow and context, using human characteristics and attitudes for CAs can be desirable to foster the user's trust and increase the user experience [51] [64]. Radziwill, Benton [51] and Zamora [64] investigated quality attributes expected from chatbots. Among other attributes, humanised attributes like sensitiveness to social concerns, personable and personality traits were identified.

Efforts in the domain of conversational interfaces often improve efficiency or tackle technical issues, but the users perspective is a recent subject of research. Piccolo et al. [49] argued that the need for a more user centered research is significant and urgent. The authors pointed out that there is a recognized potential for fulfilling emotional needs or addressing sensitive topics like privacy. However, further research is needed to asses the real value and impact for the user target group. This is especially important for applications which focus on emotional aspects or aim to achieve positive change. Thus, systematic ways to grasp the positive impact and drawbacks of these applications are necessary. Therefore, various evaluation methods are listed in the last section of this chapter (see chapter 2.5).

2.3 Positive Computing & Well-being

While existing discussions of interaction design for conversational interfaces rather focus on technical advances and issues, the objective to facilitate Positive Computing for such agents is often neglected. It is recognized by research that chatbots hold potential for positive social change which can be related to well-being [21]. Positive Computing has its origins in positive psychology [18] and tries to ensure that technology actively improves people's well-being.

Rafael A. Calvo and Dorian Peters stated in their book Positive Computing: "There are psychological factors known to increase well-being and we can use these factors to inform the design of technologies" [9]. For example, the Self-Determination Theory (SDT) of Ryan and Deci can be used to achieve sustained motivation and well-being.

Følstad et al. [21] identified the need for a research agenda for chatbots related to social good. The potential of conversational interfaces to support a positive social impact is emphasized by the authors. On the other hand, a possible negative social impact can, of course, occur as well and should be identified, mitigated or avoided. There are already chatbots on the market with a beneficial social impact like the prominent Woebot, of which the purpose is mental health support[6]. Another popular example of a chatbot with beneficial social impact is the Florence-Bot, which is in the domain of health behavior change[7]. Other application domains in the scope of chatbots for social good could be education or civic participation. Although these chatbots are related to the concept of Positive Computing, the aforementioned problems persist and to the best of the author's knowledge, no personalization nor personalityzation is applied.

Deci and Ryan's framework for understanding psychological well-being [15], also known as the Self-Determination Theory (SDT), plays an important role in the domain of Positive Computing. The authors [21] linked Deci and Ryan's three psychological basic needs *autonomy*, *competence* and *relatedness* with chatbots.

Peters, Calvo and Ryan [47] convincingly convey the importance of the three basic psychological needs (BPN), defined in Deci and Ryans Self-Determination-Theory, for the design of technology. Thus, fostering *autonomy* would increase engagement, supporting *competence* would increase motivation and facilitating *relatedness*, would enhance well-being. As the authors emphasize, these statements are not just vague proposition or opinions, but empirically proven in four decades of research and are valid across different cultures. In their work, the authors developed a model for predicting and describing the impact of the BPN to engagement, motivation and well-being. This model comprises Motivation, Engagement and Thriving in User Experience (METUX). The foundations of this model are spheres of need satisfaction and specify the relevance of each basic psychological need in various environments. For example, *relatedness* may not play an important role at the interface level but may be important at the behavioral level.

In total, six different spheres of need satisfaction were identified: *technology adoption*, *interface*, *task*, *behavior*, *life* and *society* (see figure 2). The importance of this differentiation becomes clear when a psychological need is satisfied at one level but harmed at another level. For example, a video game may satisfy all three basic psychological needs from the SDT within the first three spheres (interface, task and behavior) and hence enhance engagement, motivation and well-being at the very moment, but regarding the life sphere may have the exact opposite effect.

[6] `https://woebot.io`, *last accessed 2019-12-12*
[7] `https://www.florence.chat`, *last accessed 2019-12-12*

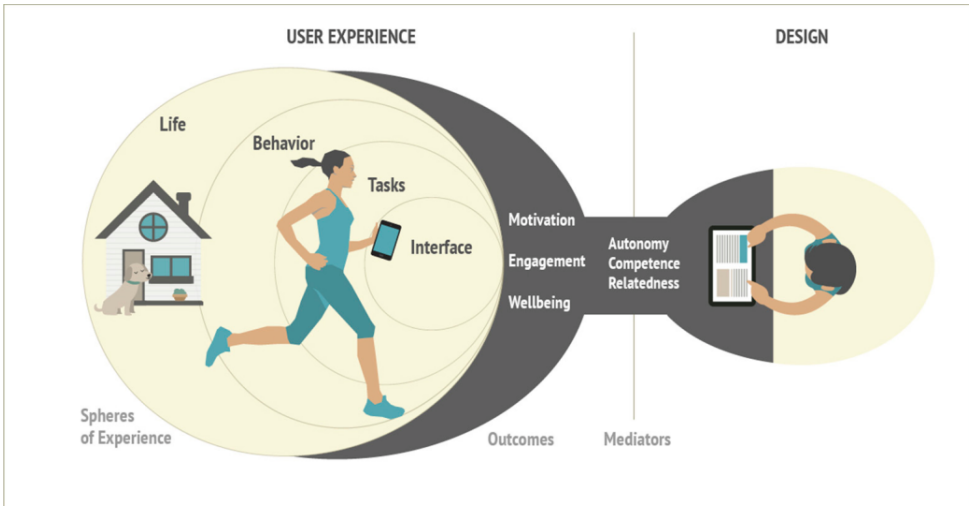


Figure 2: METUX Spheres of Psychological Need Satisfaction [47]

2.4 Personalityzation

The term personalityzation refers to an approach where UI features are adapted to an explicit model of a user's personality [4]. Personalized interfaces can utilize the user's inherent dispositions (e.g. personality traits) and improve the positive impact of technology for its users. In contrast to affective-state based personalization, this approach offers a more durable and viable way of personalizing a computer system [4]. Emotion-based interfaces are problematic in the sense that emotions are changing regularly. Therefore, adapting the user's interface based on emotions would result in an ever-changing interface and would frustrate its user. A personalityzation approach could, therefore, generate a stable model of a user's personality and adapt the interface and the behavior of the system appropriately.

Arazy et al. [4] suggest a personalityzation framework and the provision of a proof-of-concept with two empirical studies in the context of an online recommendation system. In their first study, the authors used the *emotional stability* trait to investigate the influence of a particular UI feature, namely *social anchoring*. *Social anchoring* refers to an anchoring effect where the social context and, in particular, the anchor's source elicit processes of persuasion and social influence and affect judgment [58]. This experiment has shown that participants, low on the personality trait *emotional stability*, were rather influenced by the indication of either a high respectively low rating value (*social anchor value*) of other participants. Interestingly, their ratings were influenced in the respective direction (i.e. a

higher social anchor value would lead to a higher voting of the participants). On the other hand, participants with a high level of *emotional stability* were not influenced significantly by those anchor values. In the second study the influence of different levels of the personality trait *conscientiousness*, with regard to online participation, was investigated. The experiment revealed that participants with a high level of *conscientiousness* rather submit a rating when an indication of a low critical mass value (a small number of previously added ratings) was shown. Contrarily, a high number of previously added ratings had the opposite effect on participants with a high level of *conscientiousness*. Exactly the opposite has shown to be true for attendants with a low level of the same personality trait. Hence, participants with a low level on *conscientiousness* did rather not participate when a low critical mass value was shown and tended to participate when a high number of previously submitted ratings was presented. The authors impressively convey that considering the personality of users can improve the efficiency of the user interface. Furthermore, they were able to show that empirical findings from psychology can be incorporated into the design process without the need of an intermediate transformation model. A transformation model in this context is an intermediate layer which bridges the gap between kernel theories (high level theories from natural or social science) and design theories [4].

Musitz investigated in her bachelor thesis the effects of disposition on the interface and the interaction design [35]. She listed five different disposition types concerning the user's satisfaction: *preferences*, *needs*, *goals*, *expectations* and *personality*. Considering these disposition types, personality was identified as being the most stable one. As a result of her findings, she developed ten different personas, based on the disposition of the personality, each reflecting a different interaction type. The personas were created based on the Big Five [50] and named after the predominant personality trait. The predominant traits are *outgoing*, *reserved*, *serene*, *fragile*, *controlled*, *impulsive*, *adjusted*, *reluctant*, *progressive* and *conservative*.

In her master thesis Smestad [38] provides a personality framework which follows a user centered design approach for designing a chatbot's personality. The key assumption in her thesis is that personality can provide a stable pattern on how chatbots are perceived by their users, in dependence with the user target group, the companies brand mission and the task the chatbot performs. Research has shown that personality can be used as a design variable on how and to what extent humans anthropomorphise computers. In other words, the extent of humanness determines the degree of anthromorphism. The authors furthermore emphasize the difference between anthropomorphism and humanness. Anthropomorphism in research is denoted as the psychological attribution of human features to something non-human. Humanness, on the other hand, describes to what

extent something acts or appears human. While anthropomorphism is encouraged and can provide stability in how CAs are perceived, excessive humanness can have negative results (e.g. creepy, scary appearance). This phenomenon is well known in literature and termed: "Uncanny Valley" [43], [60]. In the study of Smestad et al., two chatbot prototypes within the same domain, the same set of functionalities and the same user target group, but with different personality traits, were tested and evaluated. The modelling of the personality was again based on the Big Five. The first prototype was modelled with a high level of *agreeableness* and was the result of the user centered design personality framework. Hence, this prototype was assumed to be a good fit for the defined target group. The second prototype is designed to be the opposite of the agreeable chatbot personality and is consequently low in *agreeableness*, *extraversion* and *neuroticism* but exhibits a high level of *conscientiousness*. It was expected by the authors that this chatbot would not be an ideal fit for the defined target group. In the evaluation, the first chatbot has shown a positive effect on the user experience compared to the personality of the second chatbot. For the evaluation the AttrakDiff [28] tool was used which is a tool for assessing both the pragmatic (*efficiency*, *usefulness*, etc) and hedonic (*attractiveness*, *stimulation*) qualities of interactive interfaces.

Fromme et al. [34] conducted an interview study in the context of the individualization of a companion system. The companion system was imitated by a human actor (Wizard-of-Oz method). The ”system” asked the participants personal questions in order to individualize itself to the respective user. Additionally ”it” encouraged the participants to extensively report their experience with the system. Based on these narratives, the authors extracted four ideal types: the *non-reflecting*, the *complaisant*, the *skeptic* and the *self-doubting*. As a result of their qualitative research approach, they revealed an interesting paradox. Despite the fact that some participants ascribed negative attributes to the system, they were cooperative. The authors underlined that it was not the ”Companion Design” which guaranteed a positive user experience, but rather the individual user’s attribution to the system.

Ryan and Rigby stated that personalization creates a sense of ownership which can foster *autonomy* that again enhances engagement [47]. Individualizing a conversational agent on durable dispositions like personality has pointed a promising way to achieve sustained user engagement.

Personalized interfaces have many advantages but also risks should be mentioned and considered. As Piccolo et al. [49] mentioned, technologies which make use of personalization techniques should be treated carefully. Mensio et al. [41] warn of a ”filter bubble” where personalized interfaces hide all the opinions different from the ones expected, and as a consequence negative effects on personality traits (e.g. negative influence on the *extra-*

version personality trait) can be expected. Different negative impacts on society and it's users are discussed and identified. Addiction, isolation and decrease of empathy are among possible negative consequences. As already mentioned, addressing privacy issues is especially important for technologies which utilize personalization approaches. Users should be informed and asked for their consent, when and to what extent those techniques would be applied. This is of special importance for automated personalization methods in combination with data from social media (e.g. Facebook, Twitter, etc). Especially since the Cambridge Analytics scandal[8], where data of users were harvested without their consent, people were not pleased about the policy of Facebook. Cambridge Analytics generated user models of 270.000 Facebook users after they had answered a personality quiz. Thereafter, data of 87 million users were harvested without their knowledge[9].

2.5 Evaluation Methods

A lack of appropriated evaluation methods for conversational interfaces was identified in the preceding literature survey. Although researchers have recognized the importance of proper evaluation methods and already have developed and proposed some of them, the adoption to the field of conversational agents has just begun.

Smestad et al. used the AttrakDiff [28] method for evaluating interactive user interfaces. This evaluation method is following a user centered approach and is assessing both pragmatic and hedonic qualities. It consists of 28 opposing adjectives and ordered on a scale of intensity. Although this evaluation method also includes hedonic qualities that are opposed to conventional evaluation methods, it has not been developed for conversational interfaces. Accordingly, participants of the conducted study [38] noted that some of the adjectives did not fit to the assessment of conversational interfaces.

Bartneck et al. emphasize in their study [5] the need for a standardized measurement tool in the field of Human Robot Interaction (HRI). As a result, five key concepts from the HRI field were identified, namely: *anthropomorphism*, *animacy*, *likeability*, *perceived intelligence* and *perceived safety*. These concepts were used as a basis for five consistent questionnaires with different semantic scales (Godspeed questionnaires) [5]. The questionnaires can be used across different projects but also within the same project for different configurations. Since the field of HRI is closely related to the HCI field, an adaption of the Godspeed questionnaires seems reasonable.

[8] https://www.theguardian.com/uk-news/2018/mar/22/cambridge-analytica-scandal-the-biggest-revelations-so-far, *last accessed 2020-02-07*

[9] `https://en.m.wikipedia.org/wiki/Cambridge_Analytica`, *last accessed 2020-02-07*

Despite the awareness that conversational interfaces should encourage enjoyable conversation, the user experience literature doesn't fully cover the subject. Hence, the importance of measuring the social aspect in conversational interfaces through the focus on interpersonal communication competence (ICC) [52] was stressed by Skjuve and Brandzaeg [59]. Based on previous knowledge regarding the ICC, they modified Rubin and Martin's interpretation of ICC in order to fit in the context of conversational agents. As a result, a total of ten communication competences were used to assess the user experience:

1. *self-disclosure*
2. *empathy*
3. *social relaxation*
4. *interaction management*
5. *assertiveness*
6. *altercentrism*
7. *expressiveness*
8. *supportiveness*
9. *immediacy*
10. *environmental control*

Each of those items are represented on a Likert scale, ranging from one to five. Especially items measuring the competence "assertiveness" are important in order to determine if and to which extent the given CA is able to mitigate inappropriate behavior [59]. An extreme case and prominent example of an abusive usage of a text-generative AI based chatbot is Tay[10]. Only after a few hours online on Twitter, Tay became sexist and racist due to the abusive usage of some users. Chatbots have been targets of harassment multiple times in the past and the importance of overcoming this issue will be crucial, particularly for AI based CAs.

In the above described METUX model (see chapter 2.3), each of the evaluation methods that were dedicated to the defined spheres, including the additional sphere of the adoption of technology, were utilized. The aim of these evaluation methods is to measure need satisfaction related to the design of technology and thus enable designers to iteratively improve the outcome of the technology in question. The adoption of technology can be measured with the Self-Regulation Questionnaire (SRQ) which identifies

[10] `https://en.wikipedia.org/wiki/Tay_(bot)`, *last accessed 2019-12-12*

types of motivation (*autonomous* vs. *controlled*) with regard to the regulation of a specific behavior [53]. The authors argued that the adoption of technology is mediated by anticipated competence and thus a modification of the Perceived Competence Scale (PCS) is reasonable. As a result of these modifications, the authors developed the Autonomy and Competence in Technology Adoption (ACTA). For the assessment of the outcomes regarding the interface sphere, the Player Experience of Need Satisfaction (PENS) was adopted [54]. Peters, Calvo and Ryan named this adoption Technology based Experience of Need Satisfaction (TENS-Interface). At the interface level, this evaluation method can be used to asses the usability of the interface and the respective need satisfaction during use. The adoption itself can be again adopted to asses outcomes of the task sphere (TENS-Task). At the behavioral level, the authors suggest The Psychological Need Satisfaction in Exercise Scale (PNSES) [48] which measures perceived need satisfaction when doing exercises. The authors pointed out that in many cases there may not be a suitable measurement tool for the behavior in question. Thus, the authors recommend to use the general Basic Psychological Needs Satisfaction (BPNS) questionnaire [11] or the previously mentioned SRQ questionnaire. Last but not least an adoption of TENS for assessing the outcome regarding the life sphere was suggested by the authors (TENS-LIFE).

3 Methodology

In this chapter, the methodology used to validate the concept of this work and to answer the formulated research questions are explained.

3.1 Approach

Prior to this thesis, a research survey was conducted that has revealed a research gap in the area of chatbot personalityzation as well as in the application of the PC paradigm to CAs. A predominant focus on functionalities and features rather than on a positive user experience was identified.

In this study, a conceptually validating approach has been applied based on a mixed-method approach. Both quantitative and qualitative research methods were utilized to evaluate the concept of this paper. Based on an existing prototype, the author of this work has revised and reimplemented an existing prototype as part of his employment at the University of Vienna. The concept of the present health care chatbot prototype [61] was evaluated, extended and iteratively improved. A preliminary user test revealed a user dissatisfaction with the existing prototype. The chatbot was not adapting to its users, and the goal concept was very rudimentary. Furthermore, the logging of the user's intake was reported to work poorly. Therefore, the service responsible for the intake logging was replaced. Additionally, theories from the field of psychology were incorporated into the service design of this chatbot. A technical framework was developed to enable a disposition sensible design. Subsequently, a DSD was integrated into the service design of the present prototype, thus following an personalityzation approach. The result is an improved concept implementing the Positive Computing paradigm into the prototype.

After the development phase, a participative user study was conducted to evaluate the concept. The METUX model was utilized to structure and evaluate this project. The study started with a pre-test followed by a practical test phase and concluded with a post-test. The participants received two emails with the respective questionnaires for the pre-test as well as for the post-test. A personality test was conducted at the beginning of the practical phase to enable the DSD. Last but not least, a semi-structured interview was conducted to gain in-depth insights. The second research question was answered with a proof-of-concept. Details and results of the study design can be found in the evaluation chapter (see chapter 6). Figure 3 depicts the methodology used to answer the research questions.

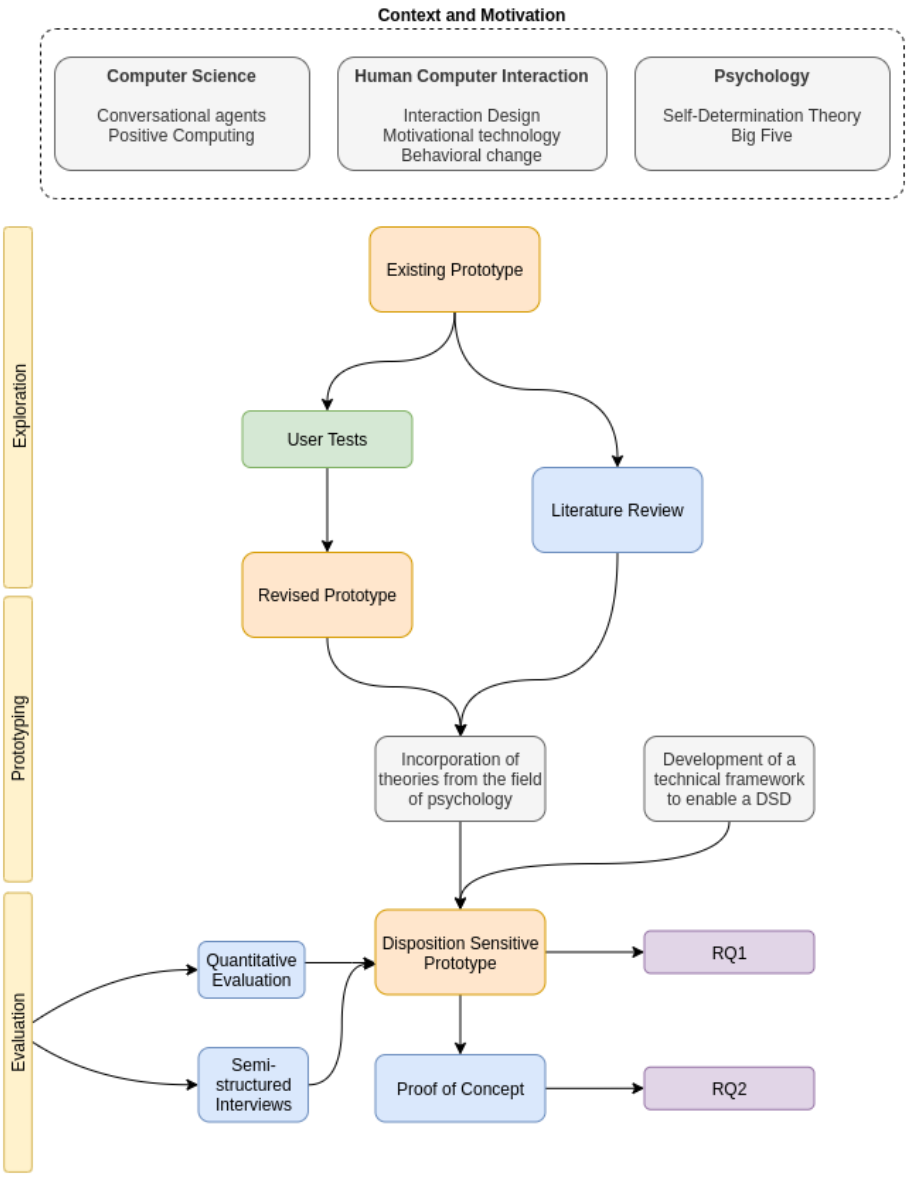


Figure 3: Outline of the methodology

4 Motivation, Engagement & Well-being

For this study a technology was designed, utilizing the Positive Computing paradigm to promote motivation and well-being through goal setting.

4.1 Design to Motivate

Motivation is a key factor for taking any kind of positive action. It is a trigger to start an activity and by sustaining this activity by an ongoing urge to carry on, we are engaged. Motivation is recognized by researchers as an important contributor to psychological well-being [9]. This is especially the case for health care technologies. Increased motivation to use health care technologies ultimately fosters well-being. This leads us to the question: How can technology be designed to foster motivation?

> *"Don't ask how we can motivate people. That's the wrong question. Ask how we can provide the conditions within which people can motivate themselves"* - Edward Deci

As this quote by Edward Deci conveys, it is not about motivating the users of technology but rather to provide the right conditions for self-motivation. At this point, the Self-Determination Theory comes into play which basically defines three psychological needs to ensure people's self-motivation and psychological well-being. The three BPNs are intrinsically rewarding and a safe target for design since the extent of need satisfaction of the BPNs cannot exceed [47]. In the METUX model, the BPNs are portrayed as mediators for motivation, engagement and well-being. The BPNs are inherited to all people but the extent of need satisfaction may vary. Therefore, a personality model is utilized to better understand the users and adaptively respond to the diverse extents of need desire (i.e. individual differences). The ability of motivation to contribute to well-being depends heavily on how autonomous it is. The concept of this work promotes, therefore, autonomous decision making through goal setting, e.g. increasing the intrinsic motivation to get involved (remember: according to CET, intrinsic motivation can be facilitated or undermined). Additionally, positive feedback is provided to enhance free-choice behavior and interest. Promoting connectedness, especially for users high on the personality trait *extraversion*, and on the other hand, avoiding social pressure for users being less extraverted, increases motivation and engagement. Increasing the perceived competence of the users' activity supports engagement which is realized through the goal engine of the conversational agent prototype.

4.2 Goal Setting

A main component of the conception of this work is the setting of goals. For this reason, findings from the goal setting theory are incorporated into the concept of this study. The goal setting theory provides a framework to investigate how defining goals influences performance, self-efficacy and satisfaction. The explicit setting of specific and difficult goals has shown that the performance of individuals and teams increased. According to the theory, both, goal progress and goal importance, are strong predictors of feelings of success and well-being [37]. Two key aspects of goal setting are feedback and commitment to the goal. Therefore, both aspects are included in the goal setting component of this study.

At the beginning of the interaction, a goal is negotiated with the user. The user can choose himself which goal he wants to pursue. A user can choose one of four different goals which are either related to physical activities or to nutritional behaviour. These goals are:

- be more active
- eat healthier
- gain weight
- lose weight

In figure 4 the conversational flow of the goal setting is illustrated.

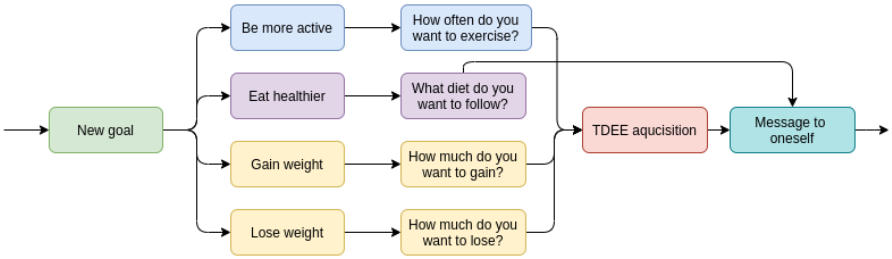


Figure 4: Goal setting

During the goal negotiation, the user can define the difficulty (e.g. number of exercise days for an activity goal or the number of kilograms he wants to lose or gain in case he picked a nutrition goal). A user can also decide to follow a diet (e.g. low-carb diet, vegan diet, etc).

4.3 Well-being

Positive Computing is concerned with achieving well-being for the users of technology, but how is well-being defined? There is no clear answer to that, as it depends on the perspective. The Positive Computing framework provides three different viewpoints to answer this question. Each of those models is supported by substantial empirical evidence and together they form the foundations of modern theories of well-being [9].

The Medical Perspective From the Medical Perspective, well-being is defined as the absence of dysfunction. To be useful for well-being supportive promotion additionally to medical treatment, this perspective appears to be too narrow.

The Hedonic Perspective The Hedonic Model focuses on the experience of positive emotion. Well-being is seen as the aggregation of pleasurable experiences. This perspective considers the subjective well-being (SWB) of people and states well-being as fundamental for maximising pleasure and avoiding negative moments. The hedonic model consists of three components: *life satisfaction*, *presence of positive mood* and *absence of negative mood*. It respects people's individual ideas on what well-being means for them (e.g. through self-reports, diaries etc.).

The Eudaimonic Perspective The Eudaimonic Perspective differentiates the concept of well-being from definitions of happiness. It can be seen as a level on which one finds meaning in life and fulfills one's great potential [9]. This perspective sees SWB research as a too narrow and flawed indicator of healthy living. Conversely to the hedonic perspective, this perspective considers the focus on fleeting emotions and the neglection of long-term overall stability as problematic. The eudaimonic model measures how well an individual performs on a set of factors that foster well-being (e.g. the basic psychological needs defined by the SDT).

None of those perspectives is right or wrong as it heavily depends on the project in question. "The PC framework is designed to let practitioners choose the suited theories; not prescribing the use of a specific theory" [9]. It rather helps us to structure the way we think about the design of technology. For this master thesis project, the main perspective used is the eudaimonic viewpoint, though the other models are considered to be also important and part of this work.

4.4 Personality Model

In order to construct a chatbot personality, a suitable personality model is needed. The most widely known and accepted framework to model personalities is the Five Factor Model (FFM), also known as the Big Five or the O.C.E.A.N (Openness to experience, Conscientiousness, Extraversion, Agreeableness, Neuroticism) model [26], [38]. The FFM is the result of decades of personality research and has its foundations in Raymond Cattell's 16 fundamental factors of personality. It is based on lexical theories, which use adjectives in natural language describing differences that are most salient and socially relevant among individuals. The validity of this model was confirmed by McCrae & Costa, two highly respected psychologists [40]. Ackerman claimed that this model can be applied in various countries and cultures and that its assessment scales are reliable and valid for measuring the five factors [2].

The theory states that five factors or dimensions are sufficient to describe the personality of human beings. In this model, each dimension has two opposing sides. Within each dimension, there are many traits and characteristics describing the respective dimension. For example, the personality trait *extraversion* and its opposite trait *introversion* are considered to form one dimension. The five dimensions of the FFM are:

1. *extraversion* vs. *introversion*
2. *agreeableness* vs. *antagonism*
3. *conscientiousness* vs. *impulsiveness*
4. *emotional instability* (*neuroticism*) vs. *emotional stability*
5. *abstract thinking* (*openness to experience*) vs. *concrete thinking*

4.5 Acquisition of Personality

A disposition sensitive conversational agent needs to understand its users and should appear and act in accordance with the personality of the respective user. Having a suited personality model, the acquisition of personality can take place. The acquisition of personality can be accomplished in various ways. Information about personality ca be acquired automatically or with the traditional approach, using self-reported assessment tools like questionnaires.

Finnerty et al. made a differentiation based on the source of data. The authors analysed automated methods to acquire personality information based on text, audio, video, mobile phones and wearable sensors [20]. For audio-based personality acquisition approaches, researchers revealed that

the *extraversion* trait is associated with a higher pitch and a higher variation of the fundamental frequency, with fewer and shorter silent phases [55], [56]. Research on communication styles revealed that extraverted people in general speak more, faster and with fewer pauses and hesitations [23]. Mohammadi et al. [42] used pitch, formants, energy and speaking rate to infer personality from speech samples. The results indicate an acceptable recognition rate for all traits except for *openness to experience*. Finnerty et al. suggest that these findings should be treated as preliminary because only the recognition rate for the traits *extraversion* and *conscientiousness* was satisfactory.

Holtgraves [30] has investigated the link between texting styles and personality traits. His research has revealed that the personality trait *extraversion* was associated with more frequent texting and the use of personal pronouns. Furthermore, extraverts have been associated with using more letters in words to add emphasis to a word (word expansions). The usage of swear words has shown to correlate negatively with participants with a high score on *agreeableness*. Schwartz et al. [45] have given evidence that demographic information and likes on Facebook can predict private attributes and traits. Various studies on Facebook data have revealed that the trait *openness to experience* is easily predictable [36] [25]. Kosinski et al. have successfully used Facebook likes and activities to infer the personality of users. [36]..

The available data on social media platforms like Facebook or Twitter has evidenced to be representative for our identity, attitudes and beliefs and various studies gave evidence that personality can be predicted to some extent by using online data [20].

Computational methods for personality acquisition seem to be a promising way to automatically infer personality for conventional Conversational Interfaces (CUIs) but also for Voice User Interfaces (VUIs) without the need of answering a long self-reported questionnaire. Despite the benefits of using automatic methods to infer personality, the implementation of such a system is out of the scope of this work. Therefore, a traditional standardized self-reported questionnaire from the International Personality Item Pool (IPIP) was used for this project. The 50-item IPIP version of the Big Five Markers was presented to the participants of the accompanying study at the beginning of the practical phase (described in the evaluation chapter 6).

4.6 Personality Match

Besides acquiring the personality of the chatbot users, the personality of the chatbot also needs to be considered. The personality of the chatbot is modelled using the Five Factor Model. Instead of having only one personality for the chatbot, the concept of this project provides a multi-personality

approach. For this purpose, five different chatbot personas have been created and integrated into the service design of this project. The personas are inspired by the Musitz personas mentioned earlier (see chapter 2.4). The results of the self-reported assessments scale can be interpreted as vectors. For example, a vector of one self-reported assessment might be looking like:

$$\vec{V}_{O.C.E.A.N} = \begin{pmatrix} 40 \\ 10 \\ 40 \\ 25 \\ 15 \end{pmatrix}$$

Each row represents one personality trait. By utilizing the vector space model, distances between individual O.C.E.A.N. vectors can be calculated. The Euclidean distance is used as the distance calculation function for the personalities of the chatbot prototype developed in the present framework. Based on the similarity attraction principle [3], [1], the chatbot personality with the smallest distance to the respective user is selected. The similarity attraction principle states that people are attracted to personalities that are similar to themselves. Musitz suggested in her design implications to use the similarity attraction principle in everyday's interactions of a DSD technology [35]. In challenging situations, she recommends to apply the complementary attraction principle which implies that people prefer personalities that are complementary to their own personality [3]. Following her suggestion to consider these principles for personalized agents, the similarity attraction principle is part of the concept of the conversational agent developed for this study.

The aim of adapting the personality of the chatbot to the user is to increase the user satisfaction and subsequently improve the overall user experience.

4.7 Disposition Sensitive Design

Each persona distinguishes itself not only by lexical differences but also regarding behavior, set of features, style of information presentation and texting frequency.

Research has revealed that people with a high score on the *extraversion* trait prefer a flat presentation of information [32]. This is argued with the higher sensory threshold of extraverts, which means that a stimulus needs a higher intensity level to be detected, compared for example with introverts. According to the finding of Kim et al., the presentation of more data to extraverts overcomes the higher sensory threshold [32], whereas introverts prefer a gradual, hierarchical structure of the presentation of information. As a result of these findings, two different versions to display information have been designed (see chapter 5).

Extraverts should be motivated by providing appropriated incentives like

the possibility to get to know new people, contrarily to introverts where social pressure should be avoided. Musitz stated that it is often sufficient to suggest a social interaction to increase the motivation for extraverts. Therefore, a social interaction feature is part of the service design of the CA and is exclusively enabled for users high on the *extraversion* personality trait.

People high on the *neurotic* trait and low on the *conscientiousness* trait tend to have difficulties with delayed gratification [24]. Therefore, the concept of this study devotes special attention to these users. In order to show intermediate results and motivate them, the developed CA prototype sends status reports and motivational quotes more frequently to its users, compared to users with a low score of these traits [31]. Additionally, the way in which success or failure is presented is different for these latter users. Research gave evidence that negative reinforcement should be avoided for neurotic people as it correlates negatively with the success of persuasion strategies [27]. Hence, in the case of a failed goal, neurotic users receive motivating sentences and are encouraged to stay on track.

Conscientious people like to surpass their own performance and like to be efficient. An emphasis on efficiency to achieve a defined goal is considered to be especially appealing to users high on *conscientiousness* [29]. To increase the perceived efficiency for high *conscientiousness* users, the CA suggests using a calendar feature in various situations. Additionally, success and failure or intermediate results are handled differently to emphasize achievements and efficiency and further motivate them.

4.8 Summary

The concept introduced in this chapter provides the conditions to facilitate the intrinsic motivation of the users. The goal setting theory is utilized to promote free decisions, thus supporting the users in their need for autonomy and competence. In order to better understand the users and their respective needs, the Big Five personality model is deployed. The personality of the user is acquired with a traditional self-regulation questionnaire and compared to the various personalities of the CA. Subsequently, the best match, with regard to the Euclidean distance, is selected and used as chatbot personality for the respective user. This idea is based on the similarity attraction principle mentioned earlier. It is assumed that the use of this particular DSD will augment the PC paradigm.

In this chapter, the concept of this study was introduced. The next chapter reveals how this concept was implemented.

5 Chatbot

The chatbot applied to test and evaluate the concept of this work is named Sam. It is a retrieval based conversational agent and it aims to support its conversational partners. Although Sam's personality is dependent on the respective user, his basic appearance stays the same. Sam's avatar is purposely a comic style human to encourage natural conversation. A realistic human image could cause negative effects, as pointed out in the related work chapter of this paper (see chapter 2). The human-like appearance facilitates the attribution of human features to the CA (anthropomorphism) without the risk of getting into the *Uncanny Valley*. Furthermore, humans need to know when they are chatting with a chatbot (and not with a human) to avoid wrong expectations.

Figure 5: Chatbot Avatar[11]

Sam is male because research gave evidence that male CAs are significantly less often confronted to negative or ambiguous language. Female chatbots are more likely exposed to negative stereotypes, sexual conversations and bad language [7]. Thus, the gender choice of an CA is an important design decision that has an impact on the user experience. Since this prototype follows a retrieval based approach rather than an generative approach, exploits or harassment are not truly feasible. Nevertheless, users trying to insult Sam will receive a warning.

[11] Creative Common licensed image

5.1 Features

This CA prototype is situated in the context of healthy living and encompasses various functionalities:

- goal engine
- fitness and nutrition goals
- recipes
- exercise recommendation
- intake logging
- logging reminder
- motivational quotes
- daily reports
- scheduling
- small talk
- message to oneself

These functionalities are partly realized through third party services (see implementation section 5.3). Some of the features and functionalities are depending on the active personality of Sam. The active personality is selected based on the O.C.E.A.N. vector of the current user (i.e. the results of the personality test). The presentation style of information (e.g. for daily reports, goal success/failure) varies also in dependence of the active personality of the CA. This is either a tabular style webview or a simple text message. The logging reminder makes sure that the user doesn't forget to log his activities and food intake. Besides the logging reminder, a user can write a message to himself. The message to oneself is a short message that can be written by a user to his future self at the beginning of the achievement of his goal. This message will be sent to the respective user close to the end of his goal. It may be a reminder of the reason for having set this specific goal in the beginning of the exercise. Such a message can have a high impact on the motivation of the user.

Special attention has been paid to error handling, as research gave evidence that user of chatbots prefer specific error messages over generic error messages. For each intent a specific error message is determined. This messages can be modified at any time without changing the source code.

5.2 Goal Setting

At the heart of Sam is the goal engine. The user can choose one of four different goals:

- be more active
- eat healthier
- gain weight
- lose weight

The length of time to achieve the goal is independent of the selected goal and is fixed to one week. After one has decided on a goal, the CA asks for relevant information on that goal. For instance, when choosing the goal to lose or gain weight, the current weight, height and age are needed to calculate the Total Daily Energy Expenditure (TDEE).

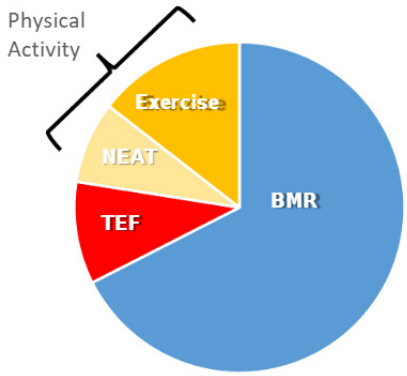


Figure 6: Total Daily Energy Expenditure (TDEE)[12]

The TDEE is an estimator of how many calories one burns a day. It consists of the Basal Metabolic Rate (BMR), the Non-Exercise Activity Thermogenesis (NEAT), the Thermic Effect of Food (TEF) and the Exercise Activity Thermogenesis (EAT) [57]. The BMR determines the estimated amount of burned calories during resting, while the EAT is an estimator for how many calories are burned during exercising. The NEAT determines the estimated amount of burned calories during non exercise activities like walking, standing or doing household chores. The TEF is the amount of calories burned during eating and digesting food. The approximate distribution is shown in figure 6. Depending on the respective goal, the user receives goal progress reports. The frequency of those reports depends on the personality of the

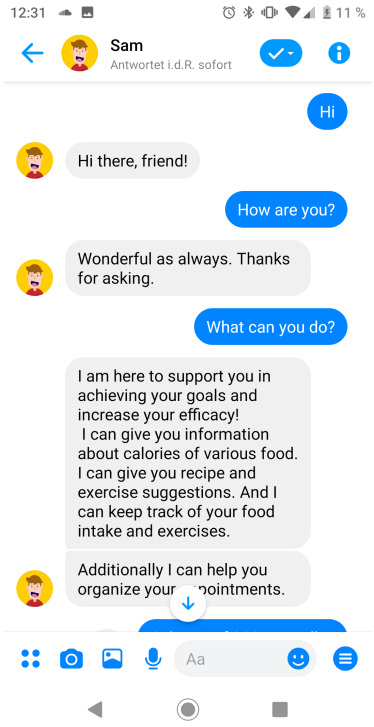
[12] `https://tdeecalculator.org/`, *last accessed 2020-02-07*

current user (e. g. a user high on the *neurotic* trait receives daily reports and motivational quotes more frequently).

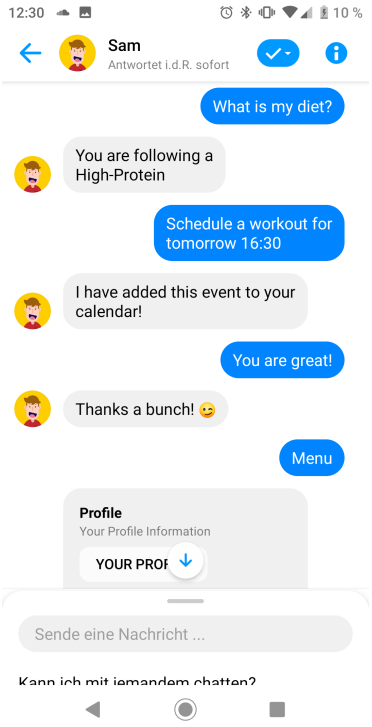
The goal engine is following the principles of S.M.A.R.T. goals [17]. S.M.A.R.T. is an acronym for: Specific, Measurable, Achievable, Realistic/Relevant, Timely/Tangible. This means, a concrete goal is much more likely to be achieved than a general goal. A specific goal defines requirements and constraints. It must establish a time frame and must be measurable. Likewise, it must be attainable, i.e. you should have an idea which steps need to be taken to achieve your goal. In the goal achievement process you develop the attitudes, abilities and skills to reach it. A realistic and relevant goal must represent an objective which you are both willing and able to work for. A goal should be set on a high level since a high goal is in many cases more likely to be accomplished than one on a lower level. The reason for it is that a low level goal exerts low motivational force[13]. Last but nor least, a S.M.A.R.T. goal must be timely and tangible, i.e. every goal needs a target date and should be within reach.

Figure 7 depicts various screenshots of a conversations with Sam on Facebook Messenger on Android.

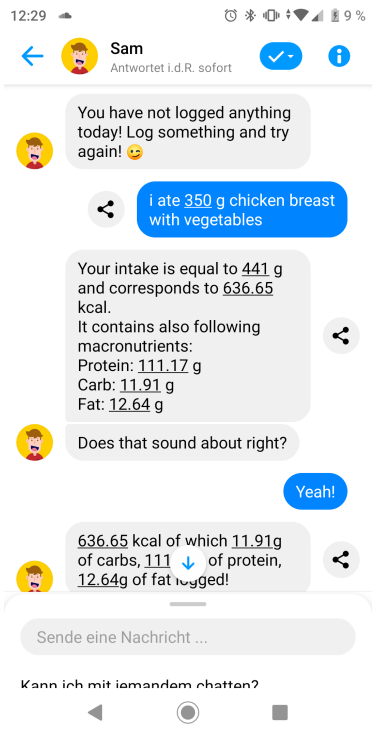
[13] `http://topachievement.com/smart.html`, *last accessed 2020-02-07*



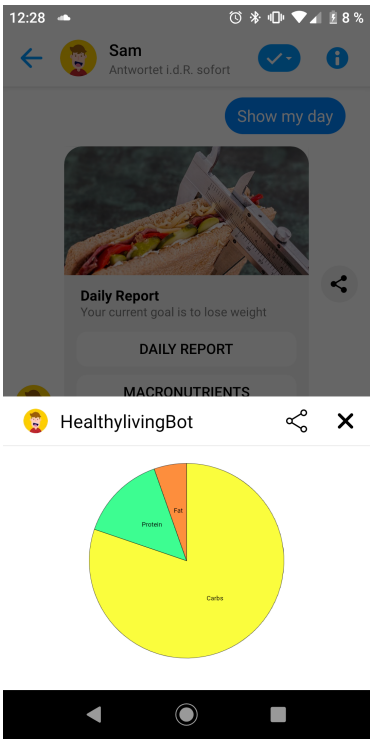
(a) Showcase of a conversation



(b) Showcase of a conversation



(c) Showcase food logging



(d) Showcase of macro nutrition visualization

Figure 7: Screenshots of Sam, the CA developed for this study.

5.3 Implementation

The main components of the implementation are the Python backend and the NodeJS Facebook wrapper. The Python backend is a RESTful API, implemented with Python 3, providing access to all third party services for the Facebook wrapper. The author of this paper has developed the backend during his employment at the University of Vienna and in the course of the present work refined and further improved. The implementation encompasses various third party services: Google Fitness, Google Calendar, Wger[14], Edamam[15], Nutritionix[16] and DialogFlow[17]. The fitness and calendar features are implemented by utilizing the Google Fitness and Google Calendar API. Most of the authentication is done with OAuth 2 standard. This approach guarantees the user control over his data by utilizing access tokens. The data generated through this CA is associated with the Google account of the user, thus not stored on the server used for this project. The usage of natural language is enabled through DialogFlow, a service offered by Google. The intake calculation is realized by utilizing Nutritionix and Edamam. Exercise recommendations are taken from Wger.

Architecture

The overall architecture is depicted in figure 8. Both, the Facebook bot and the Python backend are containerized with Docker. All external communication is TLS/SSL encrypted.

As already mentioned, the backend takes care of all the external communication and implements therefore all services aforementioned. This design guarantees modularity and enables the option to develop multiple platform independent bots without reimplementing already existing functionalities. Furthermore, a dedicated Javascript library, implementing the endpoints of the Python backend, was developed (i.e. the HLB API library and the HLB Interface).

[14] `https://wger.de/`, *last accessed 2020-03-23*
[15] `https://www.edamam.com/`, *last accessed 2020-03-23*
[16] `https://www.nutritionix.com/`, *last accessed 2020-03-23*
[17] `https://dialogflow.com/`, *last accessed 2020-03-23*

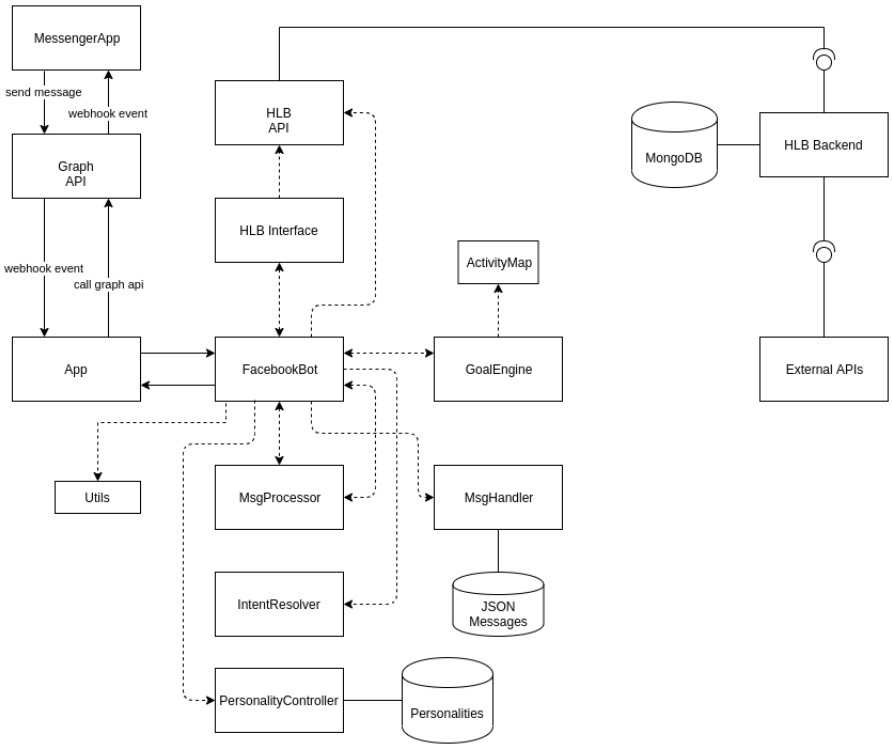


Figure 8: Healthy Living Bot Architecture

5.4 Personalities

In this section, the technical realization of the disposition sensitive design is explained. The following illustration showcases how the different personalities of the developed chatbot are implemented. In order to keep the overview, only one personality is shown in detail. The attributes and intents of the depicted personality have been reduced to a high extent because of the aforementioned reason. Thus, only representative attributes and features were kept. The personalities were realized by utilizing the JSON format. Hence, personalities can be added, deleted or modified any time, without the need of changing the source code or touching the implementation.

```
[
    "HC_LN_HA": {
        "alias": "Dorothy",
        "description":  "Organized, tendency to display self-discipline, act dutifully, and strive for achievement against measures or outside expectations.",
        "features":     "Task-oriented and goal-oriented and like to have control over things, especially their own impulses. Emphasize on efficiency.",
        "traits":       "High on Conscientiousness and Agreeableness and Low on Neuroticism (Emotionally stable)",
        "intents": {
            "greetingCommit":[
                "With that attitude you will accomplish your goals!",
                "Let's start by setting a concrete goal!"
            ],
            "weeklyRecipes": [
                "I can support you by setting a list of daily recipes onto your calendar, interested?",
                "Recipes will be added to your calendar in a bit. Have a great week!"
            ],
            "foodLoggingSuccess": [
                "Ok! KCAL_TOTAL kcal of which CARBS g of carbs, PROTEINS g of proteins, FAT g of fat logged!"
            ],
            "activityLogging": [
                "Well done! ACTIVITY_DURATION ACTIVITY_UNIT of activity logged!"
            ],
            "eatHealthierGoalAccomplished": [
                "Congratulations! You have accomplished your goal to eat healthier! ;)"
            ],
            "eatHealthierGoalFailed": [
                "You have failed your goal to eat healthier!",
                "That wasn't your best performance. You can do better than that!"
            ]
```

```
        },
        "adaptiveGoals": true,
        "useCalendar": true,
        "useMsg2Self": false,
        "appreciateSocialInteraction": true,
        "loggingReminder": {
            "appreciateLoggingReminder": true,
            "reminders":[ ]
        },
        "progressInformation": {
            "display": "tabular",
            "frequency": 3
        },
        "motivationalQuotes": {
            "quotes": [],
            "frequency": 6
        },
        "ocean" : { "O" : 25, "C" : 50, "E" : 25, "A" : 50, "N"
             : 10 }
    },
    "HO_HE_LC": {

    },
    "LO_LE_LA_HN": {

    }
    ...
]
```

Listing 1: Showcase of the technical realization of personalities in a JSON document

6 Evaluation

Based on the findings of the previously conducted literature survey, suited evaluation methods were selected and used to evaluate the concept and the prototype implementing it. In this chapter the user study itself, the participant selection and the utilized evaluation methods are explained.

6.1 Study Design

This study aimed to evaluate the concept developed for this work and the respective implementation of a health care chatbot. The study investigates how chatbots can support a healthy lifestyle and the achievement of a self-defined health goal. One main objective was to investigate if a conversational agent was able to augment or maintain the perceived competence of its users towards a healthy behaviour. The other main objective was to investigate if the METUX model could be applied to CIs and if the disposition sensible design was working in line or augmenting the Positive Computing paradigm. Subsequently, the study assessed the effects of the personalization of a chatbot on the user's behavior with regard to different health goals.

The study was structured in three phases. The first phase, the pre-test phase, consisted of two PCS questionnaires (diet improvement and exercising regularly) and two TSRQ questionnaires (eat a healthy diet, exercise regularly). After the pre-test phase the practical phase began. This phase lasted two weeks. In this phase, the participants were asked to use and test the chatbot regularly. Right at the beginning of the practical phase, the participants were asked to take a personality test. This test, a Big Five personality test, was taken from the International Psychology Item Pool (IPIP)[18]. It consisted of 50 questions which participants could answer on a five point Likert scale. The last phase was the post-test phase in which participants were again asked to fill in the same PCS questionnaires along with the System Usability Scale (SUS), TENS-Interface and TENS-Task questionnaires. The reason of asking the participants to do the PCS twice was to investigate the impact on the perceived competence after using the prototype for two weeks.

Attendees of the experiment were asked to read an instruction sheet (see appendix D) before entering the practical phase. The instruction sheet contained a short description of the chatbot and how it could be found on the Facebook platform. Furthermore, information about the various applications for accessing the Facebook platform was provided in the instruction sheet. Attendees were instructed to either use the default Facebook website (referred as desktop page), the Messenger app from the Apple Store

[18] `https://ipip.ori.org/new_ipip-50-item-scale.htm`, *last accessed 2020-02-07*

or the Google Play Store. The use of the also available Lite Messenger was discouraged, as it is limited in its functionalities and isn't able to use all features needed for this project (e.g. the carousel view). Additionally, the instruction sheet contains a listing with the instructions to get started. The participants were instructed to negotiate a goal with the CA. Which goals and which level of difficulty they would choose was left up to them. They were asked to follow further instructions from the chatbot and to regularly use it for two weeks. At the end of the first week, they were asked to set up a new goal.

Although the participants were asked to use the chatbot regularly, not all attendants have complied to that. Therefore, the frequency of the chatbot use is mentioned if relevant (i.e. very inactive or very active users are highlighted).

The study's schedule is depicted in figure 9. The grey box at the bottom of the figure illustrates the different spheres of experiences defined by the METUX model.

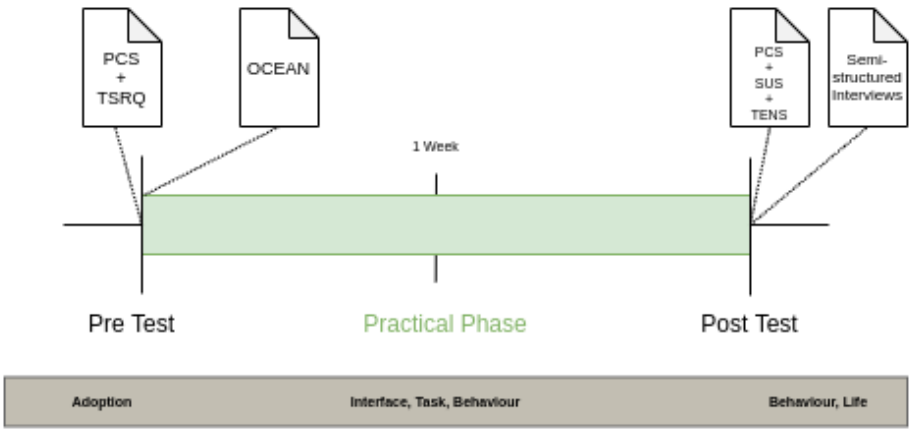


Figure 9: User Study Schedule

Participants Selection The participants of this study were recruited from the circle of acquaintances of the author. They were all between the age of 20 and 35 years and laymen regarding sport and nutrition. This target group generally has busy days and achieving fitness or nutrition goals is especially hard to accomplish for them. The developed chatbot aimed at easing the challenges for this target group. In total 16 participants were recruited, of which 10 high quality participants were left. To be considered as high

quality participant, an attendee had to have completed all questionnaires of the pre and post test phase as well as the semi-structured interviews.

6.2 Evaluating the Spheres of Experiences

The evaluation concept is partly based on the METUX model. In the course of this research project not all spheres of experiences defined by the METUX model could be considered.

Evaluating the Adoption Sphere The sphere of adoption in the setup of this research project was not evaluated since the participants have been selected by the author.

Evaluating the Interface Sphere The interface sphere was evaluated with the Technology-based Experience of Need Satisfaction (TENS-Interface) questionnaire, along with the System Usability Scale (SUS). The TENS questionnaires are based on the Player Experience of Need Satisfaction (PENS) which was developed by Ryan et al. [54]. The authors elaborated their scale on the base of the Self-Determination Theory. There are three facets of the TENS questionnaires: TENS-Interface, TENS-Task and TENS-Life. TENS assesses the degree of need satisfaction for each basic psychological need, defined by the SDT. The TENS-Interface questionnaire consists of 15 items which can be rated by the participants on a five point Likert scale. There are five items for each BPN which are all weighted equally in scoring. Regarding this evaluation, only the items concerning the BPN *autonomy* and *competence* were presented to the participants. The *relatedness* items are marked as optional by the authors. The assessment of the *relatedness* need was beyond the scope of this work.

The SUS (System Usability Scale) is a quick and reliable tool to assess the usability of a (software) system. It consists of 10 items to which participants can answer on a five point Likert scale. The final score yields a number between 0 and 100, whereas we have to be aware that these scores haven't anything to do with percentages, even though they look like. The score of a system with a good usability should be 68 or more. A system assessed with a score below 68 is considered to have serious usability issues. A score of 80.3 indicates that the usability of the system is very good. It is appreciated by the participants to the extent that they would recommend it to their friends.

Evaluating the Task Sphere The task sphere was evaluated as suggested by Peters et al. [47] with the TENS-Task scale. It consists of 12 items and participants can rate their level of agreement on a five point Likert scale. Like for TENS-Interface, all items are weighted equally and assess the need

satisfaction of the three BPN. The *relatedness* items are marked as optional by the authors. The assessment of the *relatedness* need was out of the scope of this study.

Evaluating the Behavior Sphere For this project, the behavior sphere can be evaluated using the questionnaires from the Health-Care Self-Determination Theory Packet (HCSDT)[19], as it relates to health-care behavior. The Perceived Competence Scale (PCS) was used to assess the degree to which participants felt confident about being able to make (or maintain) a change towards a healthy behavior. It is a short questionnaire, consisting of four items, which participants can answer on a seven point Likert scale. People who have a higher self perceived competence with regard to a particular behavior (practice physical exercises regularly, eat a healthy diet, do not smoking, use alcohol in a responsible way), have been found to be more likely to make and maintain a change towards positive change. The target behavior for the PCS evaluation was to eat a healthy diet and exercise regularly.

Additionally, to the PCS assessment, the Treatment Self-Regulation Questionnaire (TSRQ) was used. The TSRQ assesses the degree of a person's behavior or set of behavior being autonomous or self-determined. "The TSRQ is a set of questionnaires concerning why people do or would do some healthy behavior, enter treatment for some disease, try to change an unhealthy behavior, follow a treatment regimen, or engage in some other health-relevant behavior."[19] The TSRQ exists in different variations for healthy behavior, such as smoking cessation, diet improvement, practice physical exercises on a regular basis and drinking alcohol in a responsible way. Its consists of 15 items, six assessing the degree of autonomous motivation, six assessing the degree of controlled motivation and three assessing the degree of amotivation. Each item can be answered on a seven point Likert scale. The behavior in question for the PCS evaluation was diet improvement and exercising regularly.

Evaluating the Life Sphere The evaluation of the life sphere was outside the scope of this work. This evaluation would have required a longer time frame and multiple iterations (e.g. 3, 6 and 12 months) to show changes over a longer period, a common approach for psychology studies [47].

[19] `https://selfdeterminationtheory.org/health-care-self-determination-theory/`, *last accessed 2020-02-07*

6.3 Semi-Structured Interviews

The study was concluded with a semi-structured interview. Each interview session took about 20-30 minutes. It consisted of four main questions and ten subquestions (see appendix C). The interviews were conducted between December 2019 and January 2020. As the mother tongue of the participants was German, the interviews were conducted in German. The results of the interviews are presented in the next chapter (chapter 7).

6.4 Results

In this section the results of the quantitative research are presented. The significant results are highlighted in the following tables. To be considered as significant, at least two points in the comparison of pre and post test must have increased or decreased.

The average perceived competence of the participants for maintaining a healthier life was rated rather high (PCS Health - Pre). In comparison with the second PCS rating for maintaining a healthier life, the average rating of all participants did not change (PCS Health - Post).

PID	PCS Health - Pre	PCS Health - Post	PCS Health - Result	PCS Exercise - Pre	PCS Exercise - Post	PCS Exercise - Result
1	21	27	6	22	20	-2
2	25	24	-1	24	24	0
4	24	18	-6	22	22	0
5	28	27	-1	28	27	-1
6	26	28	2	26	28	2
8	24	26	2	27	27	0
10	24	24	0	24	24	0
11	26	24	-2	27	26	-1
12	19	20	1	21	21	0
16	22	21	-1	26	24	-2
AVG	**23,9**	**23,9**		**24,7**	**24,3**	

Figure 10: User Study PCS Evaluation

However, some participants (e.g. participant 1) showed a significantly higher PCS score. The average perceived competence of the participants to regularly exercising was rated even higher (PCS Exercise - Pre). In comparison to the second PCS rating for regularly exercising, the average rating of all participants slightly decreased (PCS Exercise - Post).

Regarding how autonomous the decision to maintain a healthier life was, the participants gave high ratings. They had a relatively low perception of how controlled their decision was to maintain a healthier life. They rated their decision little amotivated.

For the assessment of how autonomous the decision to exercise regularly was, the participants gave also a high rating, but slightly lower than for maintaining a healthier life. Similarly to maintaining a healthier life,

PID	TSRQ Diet - Auto	TSRQ Diet - Contr	TSRQ Diet - Am	TSRQ Exercise - Auto	TSRQ Exercise - Contr	TSRQ Exercise - Am
1	36	17	11	38	17	11
2	40	16	6	40	13	6
4	27	17	6	22	23	5
5	41	21	4	41	23	5
6	40	13	5	28	14	4
8	41	19	14	42	18	9
10	35	15	10	36	15	9
11	31	20	10	23	11	13
12	36	13	6	31	12	4
16	34	17	9	36	24	8
AVG	**36,1**	**16,8**	**8,1**	**33,7**	**17**	**7,4**

Figure 11: User Study TSRQ Evaluation

they had a relatively low perception of how controlled their decision was to exercise regularly. They also rated their decision little amotivated.

The interface was rated good with the TENS-Interface scale. The overall rating of the TENS-Interface is 39.7 and close to the maximum of 50. The CA also performed well on the SUS scale with 76.25 out of 100. A value of 68 is considered to be above the average and falls into category C. Only one participant gave a rating lower than 68.

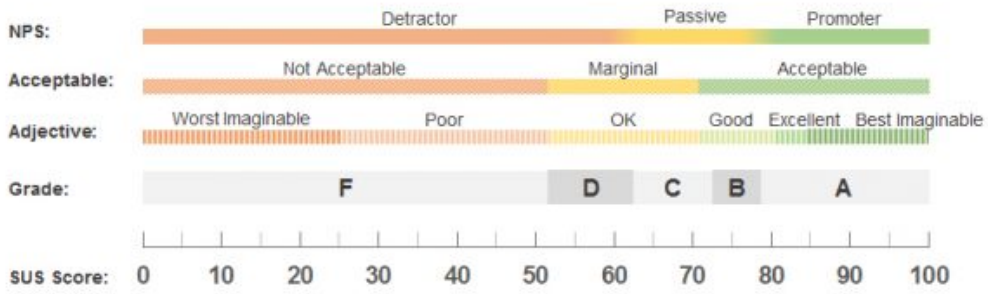


Figure 12: System Usability Scale[20]

Figure 12 depicts the ranges of the respective categories.

[20] https://measuringu.com/interpret-sus-score/, *last accessed 2020-03-31*

The TENS-Task was also rated quite good with an average value of 32.6 out of 40. The lowest rating is 29 (participant 1) while the highest rating attains the maximum value 40 (participant 6).

PID	TENS Interface	TENS Task	SUS
1	40	29	82,5
2	41	32	72,5
4	41	30	80
5	37	35	80
6	47	40	80
8	44	31	85
10	40	32	72,5
11	30	34	57,5
12	38	32	80
16	39	31	72,5
AVG	**39,7**	**32,6**	**76,25**

Figure 13: User Study Usability Evaluation

In this section, the results of the quantitative research are presented. In the next chapter, these results will be discussed and additionally, the results of the qualitative research will be revealed.

7 Discussion

As the first research question ”if a chatbot can augment/increase the perceived competence of their users in being able to make or maintain a change towards a healthy behaviour” is two-fold, it must be answered separately. The part in the sentence ”a change towards a healthy behaviour” means either to *maintain a healthier lifestyle* or to *exercise regularly*. The results of the quantitative research indicate that for three participants the perceived competence in *maintaining a healthier life* had increased significantly. For one participant the increase by 6 points was drastic (participant 1). This participant has been using the chatbot intensively each day. On the other hand, for two participants the perceived competence in *maintaining a healthier life* had decreased significantly. One other participant (participant 4) did not use the chatbot regularly. The ratings of this participant should, therefore, be treated with caution. Considering the extent to which attendees felt their decision to maintain a healthier life as autonomous, participants with an increased PCS score had also rated their motivation as rather autonomous (i.e. higher scores in the TRSQ Diet Autonomous scale). However, participants with a decreased PCS score rated their motivation to maintain a healthier life as less autonomous compared to those with a higher score in the PCS comparison. Nevertheless, this part of the first research question can be answered positively.

For the second part of the first research question (i.e. can a chatbot augment/increase the perceived competence of its user *in exercising regularly*) the result was not significant. Only one participant’s perceived competence had increased by 2 points (participant 1). On the other hand, the perceived competence of two other participants decreased by 2 points. Interestingly, attendees of which the PCS score had decreased rated their decision to exercising regularly as a rather autonomous one. Since there was only one participant of which the PCS score had increased in this category, the respective result is not very meaningful. The previously mentioned results are not very convincing neither, strictly speaking, the second part of the first research question must, therefore, be answered negatively.

The answer to the first research question is ambiguous as long as we do not consider the interaction frequency of the respective user. However, taking the frequency of use into account, the first research question can be answered with yes.

The subquestion of the first research question (to what extent is a chatbot able to increase the perceived competence to make or maintain a change towards a healthy behavior) depends on the frequency of use and on the selected goal of the respective user. For instance, participant 1 selected for both weeks a goal of the category *maintaining a healthier life*. The increase of the perceived competence for that user in this category was very strong. For the same participant, the perceived competence was decreasing by 2

points in the category of exercising regularly. Participant 4 selected a goal of the category *maintaining a healthier life*, too, but stopped using the chatbot regularly after a few days. The decrease by 6 points in that category is drastic. For this user the perceived competence in *exercising regularly* was unchanged. Participant 11 was also one of the users who didn't use the CA regularly. Hence, it can be concluded that participants using the chatbot regularly perceived the extent to which the chatbot increased their competence as strong in the category of the selected goal.

The second research question (i.e. if the Positive Computing framework/-paradigm METUX can be applied to conversational agents) can be answered affirmatively since there was no evidence of the contrary and the application of it had worked well. In the course of this work, the METUX model has proven to be very useful and structured the way of thinking in this project. However, the approach for answering this research question is a proof-of-concept and further developments are future work. The spheres of experiences defined by the METUX model help us to verify the reaching of well-being within each sphere without harming another sphere.

The subquestion of the second research question (i.e. if the disposition sensible design works in line or rather augment the Positive Computing paradigm) can be answered with yes. It worked in fact in line and augmented the Positive Computing paradigm for this project. However, as the Positive Computing framework is not prescribing any specific theories to accomplish well-being, this answer is only valid for this particular setting. As mentioned earlier, the selection of a suited theory depends on the project in question and also on the model used to define well-being (e.g. medical viewpoint). The application of the PC paradigm to a DSD has shown to be reasonable when viewed through the eudaimonic perspective. The eudaimonic model measures how well an individual performs on a set of factors that foster well-being. This set of factors represents for this project the three basic psychological needs defined by SDT. However, the extent of need satisfaction of the BPNs depends on individual differences. Hence, utilizing a disposition sensitive design for projects that are applying the PC paradigm has proven to be effective.

The results of the qualitative and quantitative research correspond to each other with regard to the perceived competence. Nevertheless, participant 10 and 12 are an exception since the results are not consistent. Participant 10 answered the question "the chatbot in being able to support his competence to accomplish his goal" with yes although the scores of the quantitative questionnaires had neither increased nor decreased in the PCS comparison. The PCS score of participant 12 had not increased significantly (increased by 1 point) although he stated that the chatbot supported his competence in achieving his goal. Additionally, attendees stated that the interaction with this CA had increased their consciousness of healthy living.

The fact that they had to log something increased their reflection and was generally positively received:

> *"Sam strengthens the awareness of what and how much one consumes"* - Participant 1[21]

The consciousness about their eating and activity habits was increased for most of the participants. One of them said that it made him want to put more variety into his diet.

Most of the participants liked the feature to log their food intake or to log their activities by sending the CA natural text (e.g. I ate a fish with 200 g of potatoes), but on the other hand, some participants had sometimes trouble to specify/enter the right data (e.g. traditional food from German-speaking participants). By some of them it was perceived as laborious that they had to google the food before they actually could log it. One of the participants who used the desktop version stated that not being able to enter things directly made him think more about it.

Through qualitative research, some interesting insights were revealed. Attendees generally appreciated the logging reminder and found the frequency with which the chatbot contacted them to be appropriate. It helped them to stay on track and it was not perceived as intrusive by anyone. One participant stated:

> *"The frequency was appropriate. Not too little so that one forgets but not annoying either."* - Participant 10

The time of the day for the logging reminder was considered appropriate especially by one of the participants:

> *"Exactly when you have already eaten at least one meal, just right and not asked too early"* - Participant 12

Some enjoyed the recipe feature while others found the proposed meals to be too exotic or complicated. The authentication process was experienced as problematic by some attendees and resulted in some functions not working properly. However, the intervention of the author mitigated this issue. The personality of the CA was not strongly perceived although it was noted as friendly and helpful. Most of the participants meant that the CA couldn't be compared to a real trainer or nutrition coach.

One participant welcomed the Message-to-oneself feature particularly:

> *"The Ms2self was nice. Why did I start? To show that in a weak moment, is a good idea"* - Participant 5

[21] the quotations have been translated into English

The motivational quotes were well received by most of the participants. One of the participants desired a more flexible and specific goal setting for the activity goal, i.e. one should not only specify the number of sessions he wants but also to set the distance he wants to manage, e.g. "I want to run 15 km this week." This participant also would have liked the possibility to specify which activities he wants to do (e.g. I want to run on Monday and do Yoga on Wednesday). One participant would have appreciated a more detailed final summary with a visualization of the goal progress (what was missing to succeed), e.g. "You have just done one workout. There are only three more to accomplish your goal!" He proposed a visualization which depicts the progress of the goal in a bar chart.

7.1 Conclusion

This research aimed to investigate the effects of applying the PC paradigm to conversational agents by utilizing a personalityzation approach. Although the importance of the user's well-being is recognized by researchers, the application of the PC framework to CAs has just begun. Various related research has revealed that personalityzation techniques can improve user experience and overall user satisfaction. Accordingly, this work, examined how personalityzation and the application of the PC can be brought together.

The quantitative and qualitative findings presented in this thesis suggest that a chatbot is at least partially able to increase the competence of its user in being able to make or maintain a change towards a healthy behavior. The quantitative outcome discussed in the previous chapter (see chapter 7) indicates ambiguous results for the first research question. Regarding the quantitative research, participants had the perception that their competence had increased after using the CA in the category of *maintaining a healthier lifestyle*. Concerning the category of *exercising regularly*, quantitative research revealed that the perceived competence of participants had decreased. In the interviews, the participants were asked directly if the CA supported them in their competence to achieve their fitness/nutrition goal. No distinction was made in the interviews between these two categories. Even though some attendees rated the increase of their perceived competence in the category of *exercising regularly* as not significant in the PCS assessment, they stated in the interviews to feel well supported in this competence. The findings reveal that the extent to which users have felt supported depends on the intensity of use of the chatbot.

The adoption of the PC framework, especially applying the METUX model to CAs, worked well in the frame of the research conducted in this thesis. It provided a way to structure and think about such a project and provided suitable evaluation methods. The evaluation methods applied in this project

proved to be feasible and well performing. Especially the PCS scale in combination with the TSRQ scales have proved to be successful. The interface was rated good as both the SUS as well as the Task-Interface evaluation have yielded good results. The DSD developed for this thesis augmented the PC paradigm effectively. The personality model used for this project helped to better understand the users and their needs, thus enabled more effective use of the PC paradigm.

7.2 Limitations

Although a total of 17 people were asked to participate in the study, only 10 of them took part in it seriously. Especially for the interpretation of the quantitative research, the number of participants was quite low. Moreover, the time frame of the study was too short to clearly demonstrate a change in behavior. Therefore, a longer-lasting study is recommended by the author (e.g. 3, 6 and 12 months). The PCS assessment should be conducted multiple times to prove a trend in the change of behavior. Furthermore, some participants had problems to follow the instructions from the instruction sheet. As a result, the authentication process or the setting of a goal was not done correctly. Additionally, the selection of the participants should have been done more strictly, since some of the participants were not using the CA regularly as it had been instructed. The process to get started was partly perceived as too complicated and thus error-prone.

The development of the CA has proven to be very laborious. The vast variety of possible user inputs and the amount of offered services were a major challenge for this project. Moreover, time constraints didn't allow further iterations (e.g. the use of an improved version of the CA). Especially the evaluation methods unfold their potential after having been applied to multiple iterations. Subsequently, comparing the results of the assessment of the different iterations, they would reveal if the changes made truly foster well-being or not.

The selected participants were all German-speaking, whereas all questionnaires and the CA were in English. Thus, especially logging the intake was perceived as difficult by some users. Furthermore, the personality of the CA could have been better perceived in the mother tongue. Since the personality test was not in their mother tongue, we had to expect some confusion possibly resulting in a biased personality model of the respective user.

7.3 Future Work

For further research, a larger-scale user study is recommended. A more comprehensive impact of using the CA can only be clearly shown with a larger number of participants, as the results of the limited quantitative research in this thesis are not very conclusive. Furthermore, to expose the effects of the personalityzation approach, a control group would be required as well. A future user study could select only English-speaking participants to avoid language difficulties. Additionally, the language model should be continuously updated and improved. In future works, the getting-started-process should be simplified, e.g. an automated personality acquisition could ease the burden of a too time-consuming getting-started-process. A future work could also develop the *relatedness* feature further and evaluate the need satisfaction of this basic psychological need. Furthermore, adaptive difficulty levels could be implemented in the goal engine to satisfy the need for competence even more.

References

[1] A. Shamekhi, M. Czerwinski, G. M. M. N., and Bennett, G. A. *An Exploratory Study Toward the Preferred Conversational Style for Compatible Virtual Agents*, vol. 10011. Springer, 2016, pp. 40–50.

[2] Ackerman, C. Big five personality traits & the 5-factor model explained. *Positve Psychology Program* (2017).

[3] Aly, A., and Tapus, A. *Towards an intelligent system for generating an adapted verbal and nonverbal combined behavior in human–robot interaction*, vol. 40 (2). Springer, 2 2016, pp. 193–209.

[4] Arazy, O., Nov, O., and Kumar, N. Personalityzation: Ui personalization, theoretical grounding in hci and design research. *AIS Transactions on Human-Computer Interaction (AIS THCI) 7* (06 2015), 43– 69.

[5] Bartneck, C., Croft, E., and Kulic, D. Measurement instruments for the anthropomorphism, animacy, likeability, perceived intelligence, and perceived safety of robots. *International Journal of Social Robotics 1*, 1 (2009), 71–81.

[6] Baumeister, R., and Leary, M. The need to belong: Desire for interpersonal attachments as a fundamental human motivation. *Psychological bulletin 117* (06 1995), 497–529.

[7] Brahnam, S., and De Angeli, A. Gender affordances of conversational agents. *Interacting with Computers 24* (05 2012), 139–153.

[8] Brandtzaeg, P., and Følstad, A. Chatbots: changing user needs and motivations. *Interactions 25* (08 2018), 38–43.

[9] Calvo, R. A., and Peters, D. *Positive Computing: Technology for Well-Being and Human Potential*. The MIT Press, 2014.

[10] Castellano-Tejedor Carmina, Moreno Jordi, C. A. P. G. L.-P. P. Preventive care infrastructure based on ubiquitous sensing (precious): A study protocol. *JMIR Res Protocols* (05 2017).

[11] Chen, B., Vansteenkiste, M., Beyers, W., Boone, L., Deci, E., van der Kaap-Deeder, J., Duriez, B., Lens, W., Matos, L., Mouratidis, T., Ryan, R., Sheldon, K., Soenens, B., Van Petegem, S., and Verstuyf, J. Basic psychological need satisfaction, need frustration, and need strength across four cultures. *Motivation and Emotion 39* (04 2014), 216–236.

[12] Cho, J. Mental models and home virtual assistants (hvas). In *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems* (New York, NY, USA, 2018), CHI EA '18, ACM, pp. SRC05:1–SRC05:6.

[13] Dale, R. The return of the chatbots. *Natural Language Engineering 22* (09 2016), 811–817.

[14] Deci, E. Effects of externally mediated rewards on intrinsic motivation. *Journal of Personality and Social Psychology 18(1)* (1971), 105–115.

[15] Deci, E., and Ryan, R. A motivational approach to self: Integration in personality. *Nebraska Symposium on Motivation. Nebraska Symposium on Motivation 38* (02 1990), 237–88.

[16] Deci, E., and Ryan, R. Self-determination theory: A macrotheory of human motivation, development, and health. *Canadian Psychology/-Psychologie Canadienne 49* (2008), 182.

[17] Doran, G. T. There's a s.m.a.r.t. way to write managements's goals and objectives. *Management Review 70* (08 1981), 35–36.

[18] E. P. Seligman, M., and Csikszentmihalyi, M. Positive psychology: An introduction. *The American psychologist 55* (02 2000), 5–14.

[19] Feinberg, S., and Murphy, M. Applying cognitive load theory to the design of web-based instruction. In *Proceedings of IEEE Professional Communication Society International Professional Communication Conference and Proceedings of the 18th Annual ACM International Conference on Computer Documentation: Technology & Teamwork* (Piscataway, NJ, USA, 2000), IPCC/SIGDOC '00, IEEE Educational Activities Department, pp. 353–360.

[20] Finnerty, A., Lepri, B., and Pianesi, F. *Acquisition of Personality*. Springer, 07 2016, pp. 81–99.

[21] Følstad, A., Brandtzaeg, P. B., Feltwell, T., Law, E. L.-C., Tscheligi, M., and Luger, E. A. Sig: Chatbots for social good. In *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems* (New York, NY, USA, 2018), CHI EA '18, ACM, pp. SIG06:1–SIG06:4.

[22] Følstad, A., B.-P. Chatbots and the new world of hci. *interactions 24*, 4 (2017), 38–42.

[23] Furnham, A. Language and personality. *Handbook of language and socialpsychology* (1990).

[24] Garcia-Argibay, M. The relationship between the big five personality traits, impulsivity, and anabolic steroid use. *Substance Use & Misuse 54*, 2 (2019), 236–246. PMID: 30375916.

[25] Golbeck, J., Robles, C., and Turner, K. Predicting personality with social media. In *CHI EA '11: CHI '11 Extended Abstracts on Human Factors in Computing Systems* (01 2011), pp. 253–262.

[26] Goldberg, L. R. *An Alternative "Description of Personality": The Big-Five Factor Structure*, vol. 59. American Psychological Association, 1990, pp. 1216–1229.

[27] Halko, S., and Kientz, J. Personality and persuasive technology: An exploratory study on health-promoting mobile applications. In *Persuasive Technology* (06 2010), Springer, pp. 150–161.

[28] Hassenzahl, M., Burmester, M., and Koller, F. Attrakdiff: ein fragebogen zur messung wahrgenommener hedonischer und pragmatischer qualität. In *Mensch & Computer 2003: Interaktion in Bewegung* (Stuttgart, 2003), G. Szwillus and J. Ziegler, Eds., B. G. Teubner, pp. 187–196.

[29] Hirsh, J., Kang, S., and Bodenhausen, G. Personalized persuasion: Tailoring persuasive appeals to recipients' personality traits. *Psychological science 23* (04 2012), 578–81.

[30] Holtgraves, T. Text messaging, personality, and the social context. *Journal of research in personality* (2011), 92–99.

[31] Höök, K. *Steps to take before intelligent user interfaces become real*, vol. 12. Institute of Electrical and Electronics Engineers (IEEE), 2 2000, pp. 409–426.

[32] J. Kim, A. L., and Ryu, H. *Personality and its effects on learning performance: Design guidelines for an adaptive e-learning system based on a user model*, vol. 43. Elsevier, 9 2013, pp. 450–461.

[33] Jain, M., Kota, R., Kumar, P., and N. Patel, S. Convey: Exploring the use of a context view for chatbots. *CHI Conference on Human Factors in Computing Systems* (04 2018), 1–6.

[34] Julia Krüger, Mathias Wahl, J. F. "es is komisch es is keen mensch" – zuschreibungen gegenüber individualisierten technischen assistenzsystemen. *Human Technology 15* (01-02 2018), 233–251.

[35] Katharina Musitz, Peter Reichel, C. L. Ux meets personalityzation: Implications for the dispositions - sensitive design of motivational technologies. *University of Vienna 55* (01 2019), 6.

[36] Kosinski, M., Stillwell, D., and Graepel, T. Private traits and attributes are predictable from digital records of human behavior. *Proceedings of the National Academy of Sciences of the United States of America 110* (03 2013).

[37] Locke, E., and Latham, G. New directions in goal-setting theory. *Current Directions in Psychological Science 15* (10 2006).

[38] Lunde Smestad, T., and Volden, F. *Chatbot Personalities Matters: Improving the User Experience of Chatbot Interfaces*. Springer, 04 2019, pp. 170–181.

[39] Löw, C. Exploring social companion technologies. In *Extended Abstracts of the 2018 NordiCHI 2018 Doctoral Consortium* (9 2018).

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

[41] Mensio, M., Rizzo, G., and Morisio, M. The rise of emotion-aware conversational agents: Threats in digital emotions. In *The Web Conference 2018* (Lyon, France, 04 2018), International World Wide Web Conferences Steering Committee, pp. 1541–1544.

[42] Mohammadi, G., V. A. M. M. The voice of personality: mapping nonverbal vocal behavior into trait attributions. In *SSPW '10: Proceedings of the 2nd international workshop on Social signal processing* (2010), vol. 8, pp. 17–20.

[43] Mori, M., MacDorman, K., and Kageki, N. The uncanny valley. *IEEE Robotics & Automation Magazine 19* (06 2012), 98–100.

[44] Oviatt, S. Human-centered design meets cognitive load theory: Designing interfaces that help people think. In *Proceedings of the 14th ACM International Conference on Multimedia* (New York, NY, USA, 2006), MM '06, ACM, pp. 871–880.

[45] Park, G., Schwartz, H., Eichstaedt, J., Kern, M., Kosinski, M., Stillwell, D., Ungar, L., and Seligman, M. Automatic personality assessment through social media language. *Journal of personality and social psychology 108* (11 2014).

[46] Peng, W., Lin, J.-H. T., Pfeiffer, K., and Winn, B. Need satisfaction supportive game features as motivational determinants: An experimental study of a self-determination theory guided exergame. *Media Psychology - MEDIA PSYCHOL 15* (05 2012), 175–196.

[47] Peters, D., Calvo, R., and Ryan, R. Designing for motivation, engagement and wellbeing in digital experience. *Frontiers in Psychology 9* (05 2018).

[48] Philip M., W., and W. Todd, R. Examining relationships between perceived psychological need satisfaction and behavioral regulations in exercise. *Journal of Applied Biobehavioral Research* (2008), 119–142.

[49] Piccolo, L., Mensio, M., and Alani, H. Chasing the chatbots: Directions for interaction and design research. In *Conversations 2018: 2nd International Workshop on Chatbot Research* (2019), Lecture Notes in Computer Science, Springer. Held in conjuntion with the 5th International conference 'Internet Science' (INSCI'2018).

[50] Raad, B. *The Big Five Personality Factors: The psycholexical approach to personality*. American Psychological Association, 01 2000.

[51] Radziwill, N., and Benton, M. Evaluating quality of chatbots and intelligent conversational agents. *Computing Research Repository (CoRR* (04 2017).

[52] Rubin, R. B., and Martin, M. M. Development of a measure of interpersonal communication competence. *Communication Research Reports 11*, 1 (1994), 33–44.

[53] Ryan, R. M., . C. J. P. Perceived locus of causality and internalization: Examining reasons for acting in two domains. *Journal of Personality and Social Psychology* (1989), 749–761.

[54] Ryan, R., Rigby, C., and Przybylski, A. The motivational pull of video games: A self-determination theory approach. *Motivation and Emotion 30* (12 2006), 344–360.

[55] Scherer, K. *Personality inference from voice quality: the loud voice of extroversion.*, vol. 8. European Journal of Social Psychology, 1978, pp. 467–487.

[56] Scherer, K. Personality markers in speech. *Cambridge University Press* (1979).

[57] Schulz, L., and Schoeller, D. A compilation of total daily energy expenditures and body weights in healthy adults. *The American journal of clinical nutrition 60*, 5 (November 1994), 676—681.

[58] Simmons, J., LeBoeuf, R., and Nelson, L. The effect of accuracy motivation on anchoring and adjustment: Do people adjust from provided anchors? *Journal of personality and social psychology 99* (12 2010), 917–32.

[59] Skjuve, M., and Brandtzaeg, P. *Measuring User Experience in Chatbots: An Approach to Interpersonal Communication Competence*. Springer, 04 2019, pp. 113–120.

[60] Skjuve, M., Haugstveit, I., Følstad, A., and Brandtzaeg, P. Help! is my chatbot falling into the uncanny valley? an empirical study of user experience in human-chatbot interaction. *Human Technology 15* (02 2019), 30–54.

[61] Suwannajng, P. Chatbot as a service composition which facilitates healthy lifestyle.

[62] Valério, F. A. M., Guimarães, T. G., Prates, R. O., and Candello, H. Here's what i can do: Chatbots' strategies to convey their features to users. In *Proceedings of the XVI Brazilian Symposium on Human Factors in Computing Systems* (New York, NY, USA, 2017), IHC 2017, ACM, pp. 28:1–28:10.

[63] Wilson, P., Rodgers, W., Blanchard, C., and Gessell, J. The relationship between psychological needs, self-determined motivation, exercise attitudes, and physical fitness. *Journal of Applied Social Psychology 33* (07 2006), 2373 – 2392.

[64] Zamora, J. I'm sorry, dave, i'm afraid i can't do that: Chatbot perception and expectations. In *HAI '17: Proceedings of the 5th International Conference on Human Agent Interaction* (10 2017), pp. 253–260.

Appendices

A Questionnaires

A.1 Perceived Competence Scales (PCS)

Scoring Information: This scale has 4 items, and an individual's score is simply the average of his or her responses on the 4 items. When the PCS is used along with the TSRQ, the PCS items have sometimes been interspersed with the TSRQ items.

A.1.1 Maintaining a Healthy Diet

Please indicate the extent to which each statement is true for you, assuming that you were intending either to permanently improve your diet now or to maintain a healthy diet. Use the following scale:

1	2	3	4	5	6	7
not at all true			somewhat true			very true

1. I feel confident in my ability to maintain a healthy diet.
2. I now feel capable of maintaining a healthy diet.
3. I am able to maintain a healthy diet permanently.
4. I am able to meet the challenge of maintaining a healthy diet.

A.1.2 Exercising Regularly

Please indicate the extent to which each statement is true for you, assuming that you were intending either to begin now a permanent regimen of exercising regularly or to permanently maintain your regular exercise regimen. Use the following scale:

1	2	3	4	5	6	7
not at all true			somewhat true			very true

1. I feel confident in my ability to exercise regularly.
2. I now feel capable of exercising regularly.
3. I am able to exercise regularly over the long term.
4. I am able to meet the challenge of exercising regularly.

A.2 Treatment Self-Regulation Questionnaire (TSRQ)

Scoring Information: This scale has 15 items: 6 that assess autonomous motivation, 6 that assess controlled motivation, and 3 that assess amotivation.

The autonomous motivation subscale consists of items # 1, 3, 6, 8, 11, & 13; the controlled motivation subscale consists of items # 2, 4, 7, 9, 12, & 14; and the amotivation subscale consists of items # 5, 10, & 15. In some previous studies the TSRQ used only 4 autonomous items, 4 controlled items, and no amotivation items.

The 8 items used in that version of the scale were: 2, 3, 4, 6, 7, 8, 13, & 14. The additional items were added to balance across subtle differences within both autonomous and controlled reasons and to allow the concept of amotivation to be examined when it is relevant to the scientific question being considered. The added items have all been validated in other self-regulation questionnaires based on SDT.

A.2.1 Diet

The following question relates to the reasons why you would either start eating a healthier diet or continue to do so. Different people have different reasons for doing that, and we want to know how true each of the following reasons is for you. All 15 response are to the same question.

Please indicate the extent to which each reason is true for you, using the following 7-point scale:

1	2	3	4	5	6	7
not at all true			somewhat			very
true			true			true

The reason I would eat a healthy diet is:

1. Because I feel that I want to take responsibility for my own health.
2. Because I would feel guilty or ashamed of myself if I did not eat a healthy diet.
3. Because I personally believe it is the best thing for my health.
4. Because others would be upset with me if I did not.
5. I really don't think about it.
6. Because I have carefully thought about it and believe it is very important for many aspects of my life.
7. Because I would feel bad about myself if I did not eat a healthy diet.

8. Because it is an important choice I really want to make.
9. Because I feel pressure from others to do so.
10. Because it is easier to do what I am told than think about it.
11. Because it is consistent with my life goals.
12. Because I want others to approve of me.
13. Because it is very important for being as healthy as possible.
14. Because I want others to see I can do it.
15. I don't really know why.

A.2.2 Exercise

The following question relates to the reasons why you would either start to exercise regularly or continue to do so. Different people have different reasons for doing that, and we want to know how true each of the following reasons is for you. All 15 response are to the one question.

Please indicate the extent to which each reason is true for you, using the following 7-point scale:

1	2	3	4	5	6	7
not at all true			somewhat			very
true			true			true

The reason I would exercise regularly is:

1. Because I feel that I want to take responsibility for my own health.
2. Because I would feel guilty or ashamed of myself if I did not exercise regularly.
3. Because I personally believe it is the best thing for my health.
4. Because others would be upset with me if I did not.
5. I really don't think about it.
6. Because I have carefully thought about it and believe it is very important for many aspects of my life.
7. Because I would feel bad about myself if I did not exercise regularly.
8. Because it is an important choice I really want to make.

9. Because I feel pressure from others to do so.
10. Because it is easier to do what I am told than think about it.
11. Because it is consistent with my life goals.
12. Because I want others to approve of me.
13. Because it is very important for being as healthy as possible.
14. Because I want others to see I can do it.
15. I don’t really know why.

A.3 Technology based Experience of Need Satisfaction (TENS)

A.3.1 Interface

Please indicate the extent to which each statement is true for you, using the following 5-point scale:

1	2	3	4	5
not at all true		somewhat		very
true		true		true

1. I feel very capable and effective at using the chatbot.
2. I feel confident in my ability to use the chatbot.
3. Learning how to use the chatbot was difficult.
4. I found the interface and controls confusing.
5. It wasn’t easy to use this chatbot.
6. The chatbot provides me with useful options and choices
7. I can get the chatbot to do the things I want it to.
8. I feel pressured by the chatbot.
9. The chatbot feels intrusive.
10. The chatbot feels controlling.

A.3.2 Task

Please indicate the extent to which each statement is true for you, using the following 5-point scale:

1	2	3	4	5
not at all true		somewhat		very
true		true		true

1. I feel confident in my ability to pursue a goal.
2. It's easy to pursue a goal.
3. I find pursuing a goal too challenging.
4. I find pursuing a gool too difficult to do regularly.
5. I feel pressured to pursue a goal.
6. I only pursue a goal because I have to.
7. I pursue a goal because other people want me to.
8. I will feel guilty if I don't pursue a goal.

A.4 System Usability Scale (SUS)

Please indicate the extent to which each statement is true for you, using the following 5-point scale:

1	2	3	4	5
not at all true		somewhat		very
true		true		true

1. I think that I would like to use this system frequently.
2. I found the system unnecessarily complex.
3. I thought the system was easy to use.
4. I think that I would need the support of a technical person to be able to use this system.
5. I found the various functions in this system were well integrated.
6. I thought there was too much inconsistency in this system.

7. I would imagine that most people would learn to use this system very quickly.
8. I found the system very cumbersome to use.
9. I felt very confident using the system.
10. I needed to learn a lot of things before I could get going with this system.

B Quantitative Research Results

PID	PCS Health - Pre	PCS Health - Post	PCS Health - Result	PCS Exercise - Pre	PCS Exercise - Post	PCS Exercise - Result	TSRQ Diet - Auto	TSRQ Diet - Contr	TSRQ Diet - Am	TSRQ Exercise - Auto	TSRQ Exercise - Contr	TSRQ Exercise - Am	TENS Interface	TENS Task	SUS
1	21	27	6	22	20	-2	36	17	11	38	17	11	40	29	82,5
2	25	24	-1	24	24	0	40	16	6	40	13	6	41	32	72,5
4	24	18	-6	22	22	0	27	17	6	22	23	5	41	30	80
5	28	27	-1	28	27	-1	41	21	4	41	23	5	37	35	80
6	26	28	2	26	28	2	40	13	5	28	14	4	47	40	80
8	24	26	2	27	27	0	41	19	14	42	18	9	44	31	85
10	24	24	0	24	24	0	35	15	10	36	15	9	40	32	72,5
11	26	24	-2	27	26	-1	31	20	10	23	11	13	30	34	57,5
12	19	20	1	21	21	0	36	13	6	31	12	4	38	32	80
16	22	21	-1	26	24	-2	34	17	9	36	24	8	39	31	72,5
AVG	**23,9**	**23,9**		**24,7**	**24,3**		**36,1**	**16,8**	**8,1**	**33,7**	**17**	**7,4**	**39,7**	**32,6**	**76,25**

Figure 14: User Study Evaluation

C Semi-Structured Interviews

Participant: 1

Hat der Chatbot dich in deiner Fähigkeit/Kompetenz, das Fitnessziel zu erreichen, gestärkt?

Warum nicht?

Wie müsste er sein, damit er deine Kompetenz (besser) unterstützt?

Sam stärkt das Bewusstsein, was und wieviel man konsumiert. Ich hatte Erfolgserlebnisse! Man fühlt sich fitter! Ich habe eine ausgeglichene Diät gewählt, weil es am einfachsten ist, dieses Ziel zu erreichen.

Mein Kompetenzgefühl wurde gesteigert. Ich kann das schaffen! Die Erinnerungen (logging reminder) helfen mir dabei, dran zu bleiben. Für Desktop Anwender: Die "Verpflichtung" etwas zu loggen, das man nicht direkt in sein Smartphone eintippt, sondern erst später, führt dazu, dass man mehr darüber nachdenkt bzw reflektiert.

Hast du das Gefühl, dass dich der Chatbot besser unterstützen kann als ein ”normaler” Trainer?

Warum kann der das besser/schlechter?

Das kann ich nicht vergleichen.

Wie hast du die Persönlichkeit des Chatbots wahrgenommen?

Hat diese zu dir gepasst?

Warst du mit den angebotenen Funktionen/Darstellungen zufrieden (Kalender, Bericht, etc.)?

Häufigkeit mit welcher der Chatbot Kontakt zu dir aufgenommen hat (Logging reminder, Motivationssprüche etc)?

Ich habe die Persönlichkeit nicht so stark wahrgenommen. Haha "manchmal nervt er mehr als sonst". Er könnte lustiger sein. Eine grafische Darstellung über meinen intake ist gut, aber es wäre hilfreich, eine Visualisierung über den eigentlichen Zielfortschritt einzubauen. Beispiel: Du hast bis jetzt eine Einheit absolviert. Dir fehlen noch x Einheiten (Visualisierung) Die Häufigkeit, in der er Kontakt zu mir aufgenommen hat, war angemessen.

Welche Eindrücke hattest du insgesamt?

Was hat dir gut gefallen?

Was hat dir nicht so gut gefallen?

Wie hast du die Interaktion im Vergleich zu einer anderen Fitness App wahrgenommen?

Hast du dir während des Interviews noch andere Dinge gedacht, die du noch anfügen/loswerden möchtest?

Ich finde es insgesamt einen coolen Ansatz! Man bekommt ein Gefühl für die Kalorien, die man konsumiert. Be more active: wenn active days konkret gesetzt sind, sollte der Chatbot auch an diesen Tagen fragen, ob ich mein Training absolviert habe.

Im Vergleich zu Runtastic: Fragen nach Größe, Gewicht und auch nach meinen Gewohnheiten zu fragen, hätte ich noch gut gefunden. Zum Beispiel: Wie viele Stunden verbringst du am Tag sitzend, stehend und in Bewegung.

Der TDEE sollte bei jedem Ziel abgefragt werden. Diese Profil Informationen habe ich nicht gefunden.

Ein ausführlicher Abschlussbericht wäre super. Zum Beispiel: Was hat gefehlt um das Ziel zu erreichen bzw. warum habe ich es nicht erreicht? Dies sollte zusätzlich visualisiert werden.

Participant: 2

Hat der Chatbot dich in deiner Fähigkeit/Kompetenz, das Fitnessziel zu erreichen, gestärkt?

Warum nicht?

Wie müsste er sein, damit er deine Kompetenz (besser) unterstützt?

Ich war oft verkühlt und konnte wenig Sport machen. Das hatte aber nichts mit dem Chatbot zum tun. Wenn ich nicht verkühlt gewesen wäre, wäre es hilfreicher gewesen. Der Chatbot hat mir irgendwelche Exercises vorgeschlagen, wenn ich meine Workouts eintragen wollte. Ich will drei mal in der Woche soviele km laufen oder ich will x mal Yoga machen. Er hat nur gefragt, ob ich etwas gemacht habe. Als er Exercises gehört hat, hat er nur irgendwelche Übungen vorgeschlagen.

Hast du das Gefühl, dass dich der Chatbot besser unterstützen kann als ein ”normaler” Trainer?

Warum kann der das besser/schlechter?

Ein normaler Trainer kann flexibler sein. Er kann Rücksicht auf meinen Zustand nehmen (z.B. wenn ich krank bin). Es könnte aber gut sein, dass der Chatbot nicht so flexibel ist. Dann hat man keine Ausreden mehr!

Wie hast du die Persönlichkeit des Chatbots wahrgenommen?

Hat diese zu dir gepasst?

Warst du mit den angebotenen Funktionen/Darstellungen zufrieden (Kalender, Bericht, etc.)?

Häufigkeit mit welcher der Chatbot Kontakt zu dir aufgenommen hat (Logging reminder, Motivationssprüche etc)?

Eh super! Ich habe das super gefunden, was er mir geschickt hat und wie er reagiert hat. Ich habe ihn gute Nacht gewünscht. Er mir auch! (lacht) b) War total ok. Die Rezepte haben mir gefallen. Ich konnte aber keine veganen Rezepte suchen. Ich habe keinen Bericht bekommen. War auch voll ok. Es war gar nicht aufdringlich. Ich hatte Angst gehabt, dass er mir voll auf die Nerven geht.

Welche Eindrücke hattest du insgesamt?

Was hat dir gut gefallen?

Was hat dir nicht so gut gefallen?

Wie hast du die Interaktion im Vergleich zu einer anderen Fitness App wahrgenommen?

Hast du dir während des Interviews noch andere Dinge gedacht, die du noch anfügen/loswerden möchtest?

Es hat mir im Großen und Ganzen gut gefallen! Ich war beeindruckt, wie gut das funktioniert. Es ist ja technisch auch sehr schwer umzusetzen. Auch die Wortkonstellationen so zu erkennen...

Ich war nicht so flexibel mit den Übungen. Ich hätte mir eine flexiblere und spezifischere Zielsetzung gewünscht. Ich konnte nur eine gewisse Anzahl an exercises angegeben. Zum Beispiel: zweimal in der Woche 3 km und einmal 5 km und einmal Yoga.

Participant: 4

Hat der Chatbot dich in deiner Fähigkeit/Kompetenz, das Fitnessziel zu erreichen, gestärkt?

Warum nicht?

Wie müsste er sein, damit er deine Kompetenz (besser) unterstützt?

Eher nicht. Ich habe die Kalorien immer eingetragen. Meine Gewohnheiten habe ich beibehalten. Aber mein Bewusstsein wurde gestärkt. Unter der Woche bin ich nicht so motiviert neue Gerichte zu machen/auszuprobieren. Ich bin nicht sehr flexibel unter der Woche. Am Anfang reinschreiben, was man so isst. Rezepte sollten dazu passen. Die man wirklich gut findet und umsetzbar sind. Zucker- und Fettwarnung (von der World Health Organisation). Leichte Belehrung würden mir helfen.

Hast du das Gefühl, dass dich der Chatbot besser unterstützen kann als ein ”normaler” Trainer?

Warum kann der das besser/schlechter?

Ein echter Trainer kann deutlich akkurater Aussagen machen. Der Chatbot kann aber noch verbessert werden. Es kann ziemlich gut funktionieren, ersetzt aber keinen echten Ernährungsberater.

Wie hast du die Persönlichkeit des Chatbots wahrgenommen?

Hat diese zu dir gepasst?

Warst du mit den angebotenen Funktionen/Darstellungen zufrieden (Kalender, Bericht, etc.)?

Häufigkeit mit welcher der Chatbot Kontakt zu dir aufgenommen hat (Logging reminder, Motivationssprüche etc)?

Ja, hat gepasst. Freundlich und lustig. Freundlich, aber nicht aufdringlich. Die Witze waren lustig! Die Google-Authentifizierung war problematisch. Den Report habe ich nicht bemerkt. Er hat sich einmal am Tag gemeldet, ob ich was eingetragen haben. Ich habe keine Motivationssprüche oder Berichte bekommen. Die Kalendereinträge habe ich nicht bemerkt. Ich war vielleicht mit dem falschen Account eingeloggt.

Welche Eindrücke hattest du insgesamt?

Was hat dir gut gefallen?

Was hat dir nicht so gut gefallen?

Wie hast du die Interaktion im Vergleich zu einer anderen Fitness App wahrgenommen?

Hast du dir während des Interviews noch andere Dinge gedacht, die du noch anfügen/loswerden möchtest?

Es hat eigentlich immer alles geklappt. Meine Fragen wurden beantwortet. Ich fand es ziemlich übersichtlich. Die help-Funktion ist super! Essen einzutragen, das etwas spezieller war, war problematisch. Es musste mehrfach eingegeben werden. Wenn es nicht funktioniert hat, wurden aber Rezepte vorgeschlagen.

Die Interaktion war direkter. Man tippt was ein und bekommt gleich eine Antwort. Wenn es etabliert ist, funktioniert das deutlich besser! Daten können einfacher eingetragen werden. Tipps können direkt gegeben werden. Chatbots sind die Zukunft! Ich hätte gerne mehr Zeit dafür gehabt.

Participant: 5

Hat der Chatbot dich in deiner Fähigkeit/Kompetenz, das Fitnessziel zu erreichen, gestärkt?

Warum nicht?

Wie müsste er sein, damit er deine Kompetenz (besser) unterstützt?

Die Makronährstoff- und kcalangaben waren sehr gut. Ein großes Plus. Wieviel habe ich gegessen, wieviel habe ich noch offen? Die Fitnessübungen waren sehr gut, vor allem für Anfänger. Es wurde von Anfang an eine rote Linie vorgegeben. Zum Beispiel: Drei mal Sport in der Woche machen. Die Texterkennung könnte besser sein. Es waren zu exakte Eingaben notwendig!

Eingaben wurden teilweise falsch erkannt. Wenn man die Formulierungen kennt, hat es eh gut funktioniert. Das Icon auf Android mit dem Chatbot Avatar war gut. Es hat mir gefallen. Man wird erinnert! Ich finde das sehr wichtig.

Hast du das Gefühl, dass dich der Chatbot besser unterstützen kann als ein ”normaler” Trainer?

Warum kann der das besser/schlechter?

Ein Trainer im Fitnesscenter ist nicht immer neben einem. Für komplette Anfänger beim Workouts ist ein Fitness Trainer besser. Vor allem, wenn man Übungen falsch macht. Insgesamt kann man das nicht gut vergleichen. Es ist einfach etwas was anderes.

Wie hast du die Persönlichkeit des Chatbots wahrgenommen?

Hat diese zu dir gepasst?

Der hellste ist er nicht. Die Persönlichkeit habe ich nicht so stark wahrgenommen. Er ist immer freundlich und nett.

Warst du mit den angebotenen Funktionen/Darstellungen zufrieden (Kalender, Bericht, etc.)?

Die Motivationssprüche waren gut, aber es waren zu wenige! Der Logging reminder war sehr wichtig. Man vergisst das Loggen und verliert sein Ziel mit der Zeit aus den Augen! Ich finde das war einer der positivsten Punkte!

Häufigkeit mit welcher der Chatbot Kontakt zu dir aufgenommen hat (Logging reminder, Motivationssprüche etc)?

Habe ich positiv wahrgenommen. Die Häufigkeit war angemessen, nicht penetrant. Man kann ihn auch gut weggdrücken. Tägliche Berichte waren geil, um zu zeigen, auf welchen Weg ich bin. Man wird erinnert, wo man heute steht. Tägliches Ziel wurde erreicht bzw. Ziel nicht erreicht. Ich habe nicht jeden Tag etwas eingetrangen. Das war mir zu aufwändig! Ms2self war gut. Triggerphrase. Warum habe ich angefangen? Das in einem schwachen Moment zu zeigen, finde ich sehr gut!

Welche Eindrücke hattest du insgesamt?

Was hat dir gut gefallen?

Was hat dir nicht so gut gefallen?

Wie hast du die Interaktion im Vergleich zu einer anderen Fitness App wahrgenommen?

Hast du dir während des Interviews noch andere Dinge gedacht, die du noch anfügen/loswerden möchtest?

Das Positive habe ich schon gesagt. Die Verständlichkeit von Eingaben könnte noch verbessert werden. Spracherkennung wäre auch gut für einen Fitnessbot. Sam ist nicht der Hellste, hahah. Im Wesentlichen: Man kommt in das Ganze gut rein! Die Aufteilung der Makronährwerte fand ich sehr gut. Die Kalorien und Aufteilung der Makronährwerte geht ungefähr in die richtige Richtung, stimmt ca. Der Chatbot ist nicht für Fortgeschrittene. Die Zieldauer war zu kurz. Eine Woche für Fitnessziele ist zu wenig. Ist aber für kleinere Ziele ok.

Participant: 6

Hat der Chatbot dich in deiner Fähigkeit/Kompetenz, das Fitnessziel zu erreichen, gestärkt?

Warum nicht?

Wie müsste er sein, damit er deine Kompetenz (besser) unterstützt?

Gerade am Anfang hat er das gut gemacht, als ich noch alles in der Früh eingetragen habe. Am Anfang war er penetrant. Ich hatte eine anti-Handy Phase. Ich habe drei Mal Erinnerungen am Tag bekommen. Er hat an einem Tag gar nicht geantwortet.

Um besser zu sein, sollte das Vokabular aufgebessert werden. Er hat viel nicht verstanden. Ich habe mich gewundert, dass ich am Ende der Woche keine Zusammenfassung bekommen habe. Er hätte mit mir schimpfen sollen, da ich viel zu wenig gegessen habe! Ich brauche einen Trainer, der mir einen Arschtritt gibt.

Hast du das Gefühl, dass dich der Chatbot besser unterstützen kann als ein ”normaler” Trainer?

Warum kann der das besser/schlechter?

Das ist schwierig zu sagen. Es ist nicht vergleichbar. Er ist doch ein Roboter! Die emphathische Komponente fehlt! Zum Beispiel: Wie man in

gewissen Situationen mit jemanden angemessen umgehen muss/soll.

Wie hast du die Persönlichkeit des Chatbots wahrgenommen?

Hat diese zu dir gepasst?

Die habe ich nicht so stark wahrgenommen. Am Anfang schon, als er mich erinnert hat, was ich machen soll (food loggen). Z.B. wenn du nicht ganz überzeugt bist, aufhören. Are you 100 % sure message.

Warst du mit den angebotenen Funktionen/Darstellungen zufrieden (Kalender, Bericht, etc.)?

Ich habe einmal probiert, einen Tag in den Kalender einzutragen. Er hat den Termin am falschen Tag eingetragen. Teilweise. Beim food loggen aber nicht besonders. Z.B. ist das Essen korrekt (mit Zahlen)? Bestätigen (mit Zahlen)? Dann nochmal eine Antwort mit Zahlen zu bekommen, war nicht so charmant. Es könnte weniger numerisch sein und besser dargestellt werden.

Häufigkeit mit welcher der Chatbot Kontakt zu dir aufgenommen hat (Logging reminder, Motivationssprüche etc)?

Er hat recht spät immer erinnert, das Essen und die Aktivitäten zu loggen. Er könnte schon zu Mittag das erste Mal erinnern und dann am Abend wieder.

Welche Eindrücke hattest du insgesamt?

Was hat dir gut gefallen?

Was hat dir nicht so gut gefallen?

Hast du dir während des Interviews noch andere Dinge gedacht, die du noch anfügen/loswerden möchtest?

Insgesamt war das Bild schon recht gut. Der Chatbot hat gute Funktionen. Dass er die Übungen vorzeigt und sogar animiert, fand ich toll. Die Übungen in den Ausführungen werden gezeigt! b) Das habe ich schon erwähnt.

Wie hast du die Interaktion im Vergleich zu einer anderen Fitness App wahrgenommen?

Ich finde das eher eine Geschmackssache. Garmin z.b gefällt mir. Da Dinge mehr automatisch passieren. Wenn ich will, kann ich die Daten abfragen

und anzeigen lassen. Ich bin in eher der statische Typ. Für mich persönlich ist es nicht ganz meines.

Participant: 10

Hat der Chatbot dich in deiner Fähigkeit/Kompetenz, das Fitnessziel zu erreichen, gestärkt?

Warum nicht?

Wie müsste er sein, damit er deine Kompetenz (besser) unterstützt?

Es hat großteils gut funktioniert. Teilweise hat er sich aufgehängt. Das food logging war ungenau. Die Erinnerungen waren gut! Ja, ich hatte das Gefühl gut unterstützt zu werden. Es hat mir geholfen bei der Stange zu bleiben.

Er sollte noch mehr motivieren! Zum Beispiel: Du hast schon einmal Sport diese Woche gemacht. Bleib dran!

Hast du das Gefühl, dass dich der Chatbot besser unterstützen kann als ein ”normaler” Trainer?

Warum kann der das besser/schlechter?

Ein echter Fitness Trainer ist besser. Er kann viel persönlicher auf dich eingehen! Es ist aber nicht wirklich vergleichbar. Es ist ein guter Ansatz für eine Alternative!

Wie hast du die Persönlichkeit des Chatbots wahrgenommen?

Hat diese zu dir gepasst?

Er war ganz niedlich. Ich hatte mir ein kleines Mankerl dahinter vorgestellt. Die Persönlichkeit hat gut zu mir gepasst.

Warst du mit den angebotenen Funktionen/Darstellungen zufrieden (Kalender, Bericht, etc.)?

Ich habe die Rezepte eher weniger verwendet. Ich habe mein Essen eher manuell eingegeben. Die Darstellungen haben mir gut gefallen. Z.B. bei Rezepten war das Karusell gut.

Häufigkeit mit welcher der Chatbot Kontakt zu dir aufgenommen hat (Logging reminder, Motivationssprüche etc)?

Die Häufigkeit war angemessen. Nicht zu wenig, sodass man vergisst, aber auch nicht nervend.

Welche Eindrücke hattest du insgesamt?

Was hat dir gut gefallen?

Was hat dir nicht so gut gefallen?

Wie hast du die Interaktion im Vergleich zu einer anderen Fitness App wahrgenommen?

Es war ziemlich gut gemacht, aber ausbaufähig. Hat definitiv Potential. Der Chatbot hat richtig gute Funktionen. Ich hatte Probleme mit food logging: Ich bin in einer Schleife gelandet bei food logging: Willst du das Essen loggen? Wurde mit ja beantwortet. Dieselbe Frage kam 3-4 Mal.

Hast du dir während des Interviews noch andere Dinge gedacht, die du noch anfügen/loswerden möchtest?

Das ist schwer zu vergleichen. Der Bot hat kindlicher gewirkt. Die Idee ist richtig gut, es ist was Anderes, mit wem zu chatten.

D Instruction Sheet

Evaluation of a Disposition Sensitive Health Care Chatbot

Thank you for participating in the user test! The aim of this user study is to evaluate a health care chatbot. This chatbot will help you in achieving a self defined health goal.

Requirements

- Facebook Account
- Facebook Messenger or web app (Desktop Version)
- Google Account

Procedure

- Consent
  - Information Sheet
  - Agreement
- Pre Test
  - Questionnaires
  - Personality Test
- Prototype Test
- Post Test
  - Questionnaires
  - Interview

Instructions

- At the beginning of this user test you need to make a personality test. This test contains 50 questions. The chatbot will adapt to your personality accordingly.
- Please use the chatbot regularly for a period of no less than two weeks.
- You need to set up a goal for each of the two weeks. At the end of the first week, the chatbot will ask you to set up a new goal for the second week.
- You can negotiate your individual goal with the chatbot.

- It is important that you use the chatbot on a daily basis in order to measure your goal progress (especially for health and weight goals).
- For an activity goal it is important to log all activities.
- For health and weight goals you have to log your food intake.
- At the end of the user test (after the second week) a final assessment will be conducted.

Instructions

- Use either the Facebook Webapp or the Facebook Messenger. The lite Messenger of Facebook is not recommended!
- Interacting with the chatbot is only possible in English!

Hints

- If the chatbot doesn't understand your request, try to rephrase it!
- If the chatbot doesn't know the meal you are trying to log, search for the recipe (e.g. send me a fish recipe) and hit the log this food button
- Type 'help' to get usage information.
- Type 'menu' for an short overview.
- You can get goal progress information anytime by typing: 'show my goal'.

Getting started

Search on Facebook for: @SamTheHLB

1. Say hi E.g. 'hi'
2. Start personality test E.g. 'start personality test'
3. Begin conversation with 'start'
4. Authenticate Google services by typing 'auth google'.

   Please ignore the warning: app not verified and hit the proceed link.

   This does NOT mean this app is unsafe!
5. Choose a goal by following the dialog or type 'new goal'

6. Log your activities and food intake ’i ate a fish’
   ’i ate 200 g of rice and 100 g meat’
   ’i went running for 30 min’
   ’i played basketball for 1.5 hours’

Demo

Take a look at the demo video!

E Information Sheet

Evaluation of a Disposition Sensitive Health Care Chatbot

Invitation paragraph

You are being invited to take part in a research study. Before you decide whether or not to take part, it is important for you to understand why the research is being done and what it will involve. Please take time to read the following information carefully.

What is the purpose of the study?

The aim of this user study is to evaluate a health care chatbot. In this study we investigate how chatbots can support a health lifestyle and in achieving a self defined fitness/health goal. This study will assess the effects of personalization of a chatbot to the user's behaviour with regard to different health/fitness goals. This study is structured in three phases. The total duration of this study is two weeks. The first phase, the pre test starts with questions regarding fitness/health topics followed by a personality test. The personality test incorporate 50 questions. Thereafter the practical phase begins. This phase lasts two weeks. In this phase you are being asked to use and test the chatbot regularly. To finalized the test, a final post test will be conducted which will asses the general usability and various impacts of the chatbot on your perceived competence and health behaviour.

Why have I been invited to participate?

You are invited to the test because you fit into the target group of this study. The target group consists of mainly young adults which are laymen to sport/health topics. It is planned to ask 10 to 15 people to participate to this study.

Do I have to take part?

It is up to you to decide whether or not to take part. If you do decide to take part you will be given this information sheet to keep and be asked to sign a consent form. If you decide to take part you are still free to withdraw at any time and without giving a reason.

What will happen to me if I take part?

The health care chatbot will assist you in achieving a self defined fitness/health goal. You will be asked to set up a health goal and log your activities and/or

food intake on a daily basis. Additionally you will be asked to answer the questions of the pre and post test. The fitness data collected during this test will be added to your Google Fit data (Google Account). Your fitness data is not stored in the database of this project. Therefore you will be asked to authenticate your google account at the start of the conversation with the chatbot. The data collected during this test will be anonymized. The test duration is two weeks. No physical attendance is required.

What are the possible disadvantages and risks of taking part?

There are no known risks/disadvantages to participate in the test albeit during physical activities injuries are possible.

What are the possible benefits of taking part?

Possible benefits of taking part to this study is an increased fitness and health awareness.

Will my information in this study be kept confidential?

All personal data will be held in accordance with the Data Protection Act (1998). All information related to you will be treated with the strictest of confidence and stored securely and will only be used for the purposes you have expressly permitted. You will not be identifiable as a named individual in any report or publication arising from this study.

What should I do if I want to take part?

If you agree to participate you will receive an additional sheet with detailed instructions.

What will happen to the results of the research study?

The results of this study will be used for the master thesis of the author. This thesis will be published. A digital copy can be requested via email.

Who is organising and funding the research?

This research is organised by the research group Cooperative Systems of the University of Vienna. The supervisor of this project is Univ.-Prof. Dipl.-Math. Dr.Peter Reichl and the co-supervisor is Dipl.-Ing.Christian Löw, Bsc. This research is conducted by a student as part of the master study media informatics.

Who has approved this study?

This research has been approved by the ethical review process of the research group Cooperative Systems of the University of Vienna.

Contact for Further Information

If you require any further information or have questions about taking part in this research, please contact the author.

Thank you for taking the time to participate to this study! We appreciate your honesty and willingness to assist with this research.

F Consent Agreement

Evaluation of a Disposition Sensitive Health Care Chatbot

I have read and understood the information about the 'Evaluation of a Disposition Sensitive Health Care Chatbot' technology study organised by the University of Vienna. I would like to take part.

I have read the information leaflet carefully and understand that:

This project is exploring how technology can influence behavioural change.	YES	NO
I will discuss what technologies might motivate and help with health care goals.	YES	NO
The use of the technology and views on the experience may be stored in various ways for further scientific use by the project team.	YES	NO
All information about me will be confidential and my identity will not be disclosed outside the research team.	YES	NO
I can withdraw from the case study at any time.	YES	NO

G Abstract English

This study investigates the effects of applying the Positive Computing paradigm to conversational agents by utilizing a personalityzation approach. Although the importance of the user's well-being is recognized by researchers, the application of the Positive Computing framework to conversational agents has just begun. Related research has revealed that personalization techniques can improve user experience and overall user satisfaction. Facilitating a satisfying user experience is essential, especially in the context of motivational and behavioral change technologies. This paper examines how personalization and the application of Positive Computing can be brought together. Therefore, this thesis investigates how a personalityzation approach can facilitate the Positive Computing paradigm to increase the perceived competence to make or maintain a change towards a healthy behavior. Furthermore, the research interest is to find out whether the METUX model can be applied to conversational agents.

Findings from a previously conducted research survey were used to form the concept of this work. An existing health care chatbot was analyzed and revised. As a result, a disposition sensible health care chatbot was developed to test and evaluate the concept. A conceptually validating approach was applied based on a mixed-method approach. Thus, both quantitative and qualitative research methods were utilized in a participative user study. A personality test was conducted to acquire the personality of the respective users and to better understand their specific needs. The METUX model was used to structure and evaluate this project. Finally, semi-structured interviews were conducted. The evaluation of the user study indicates that participants perceived their competence to maintain a healthier lifestyle as increased after using the chatbot. Participants with an increased perceived competence score (PCS) rated their motivation to maintain a healthier lifestyle as rather autonomous compared to participants of a decreased PCS score. Conversely, participants perceived their competence to exercise regularly as being lower after having used the chatbot. Interestingly, these attendees rated their decision as rather autonomously taken. The analysis of the semi-structured interviews reveals that participants felt supported in their competence to achieve their health goals. They appreciated the features of the conversational agent and stated that their consciousness of healthy living had increased.

The findings indicate ambiguous results for the first research question. Thus, further research will be required with the help of a large-scale user study. Nevertheless, the adoption of the PC framework worked well in this study. The personalityzation approach augmented the PC paradigm successfully.

H German Abstract

Diese Studie untersucht die Auswirkungen der Anwendung des Positiv-Computing-Paradigmas auf Gesprächsagenten, indem ein Personalisierungsansatz angewendet wird. Obwohl die Bedeutung des Wohlbefindens der Software-Benutzer von Forschern anerkannt wird, hat die Anwendung des Positiv-Computing-Frameworks auf Gesprächsagenten gerade erst begonnen. Verwandte Forschungen haben gezeigt, dass Personalisierungstechniken die Nutzererfahrung und allgemein die Benutzerzufriedenheit verbessern können. Es ist von wesentlicher Bedeutung, ein zufriedenstellendes Erlebenis zu ermöglichen, insbesondere im Zusammenhang mit Motivations-Technologien bzw. mit Technologien, die nach einer Verhaltensänderung streben. In dieser Arbeit wird untersucht, wie Personalisierung und die Anwendung des Positiv-Computing-Paradigma zusammengebracht werden können. Diese Arbeit analysiert daher, wie ein Personalisierungsansatz das Positive-Computing-Paradigma fördern kann, um die wahrgenommene Kompetenz zu erhöhen, eine Veränderung in Richtung eines gesunden Verhaltens vorzunehmen oder aufrechtzuerhalten. Darüber hinaus besteht das Forschungsinteresse darin herauszufinden, ob das METUX-Modell auf Gesprächsagenten angewendet werden kann.

Die Erkenntnisse einer zuvor durchgeführten Literaturrecherche wurden für das Konzept dieser Arbeit verwendet. Ein bereits bestehender Gesundheits-Chatbot wurde dafür analysiert und überarbeitet. Als Ergebnis wurde ein dispositionssensitiver Gesundheits-Chatbot entwickelt, um das Konzept dieser Arbeit zu testen und zu bewerten.

Dafür wurde ein konzeptuell validierender Ansatz angewandt, der auf einem gemischten Methodenansatz basiert. In einer partizipativen Nutzerstudie wurden daher sowohl quantitative als auch qualitative Forschungsmethoden verwendet. Ein Persönlichkeitstest wurde durchgeführt, um die Persönlichkeit der jeweiligen Nutzer zu erfassen und die unterschiedlichen Bedürfnisse der Nutzer besser verstehen zu können. Das METUX-Modell wurde zur Strukturierung und Evaluierung dieses Projekts eingesetzt.

Abschließend wurden semi-strukturierte Interviews durchgeführt. Die Auswertung zeigt, dass die Teilnehmer ihre Fähigkeit zur Aufrechterhaltung eines gesünderen Lebensstils nach der Benutzung des Chatbots als erhöht empfanden. Teilnehmer mit einer erhöht wahrgenommenen Kompetenzpunktezahl bewerteten ihre Motivation, einen gesünderen Lebensstil beizubehalten, als eher autonom im Vergleich zu Teilnehmern mit einer verringerten Kompetenzpunktezahl. Umgekehrt empfanden die Teilnehmer ihre Kompetenz, regelmäßig zu trainieren, nach der Benutzung des Chatbots als geringer. Interessanterweise bewerteten diese Teilnehmer ihre Entscheidung als autonom. Die Analyse der semi-strukturierten Interviews zeigt, dass sich die Teilnehmer durch den Chatbot in ihrer Kompetenz unterstützt fühlten, ihre Ziele zu erreichen. Die Teilnehmer schätzten die Ei-

genschaften des Gesprächsagenten und gaben an, dass ihr Bewusstsein für eine gesunde Lebensweise durch ihn erhöht worden war.

Die Antwort auf die erste Forschungsfrage ist aufgrund der vorliegenden Ergebnisse nicht eindeutig. Daher sind weitere Untersuchungen in einer groß angelegten Nutzerstudie erforderlich. Dennoch hat die Anwendung des Positive-Computing-Frameworks im Zuge dieser Arbeit gut funktioniert, wie die Bearbeitung der zweiten Forschungsfrage durch ein proof-of-concept gezeigt hat. Der verwendete Personalisierung-Ansatz hat das Positive-Computing-Paradigma wirkungsvoll ergänzt.