Changebots - Designing Chatbots to Support Blood Donor Behaviour Change


Helena M. Müller
Technische Universität Darmstadt


Melanie Reuter-Oppermann
Maastricht University

Changebots - Designing Chatbots to Support Blood Donor Behaviour Change

Helena Monika Müller
Technical University of Darmstadt, mueller@is.tu-darmstadt.de

Melanie Reuter-Oppermann
University of Twente, m.n.reuter-oppermann@utwente.nl

Changebots - Designing Chatbots to Support Blood Donor Behaviour Change

Completed Research Paper

Helena M. Müller
Technical University of Darmstadt
Darmstadt, Germany
mueller@is.tu-darmstadt.de

Melanie Reuter-Oppermann
University of Twente
Enschede, The Netherlands
m.n.reuter-oppermann@utwente.nl

Abstract

Even though blood products cannot be produced artificially, but are important for many surgeries and treatments, less than 1 % of the population donates blood in countries like South Africa or Ghana. Therefore, efficient and successful blood donor mobilisation and management are important. We argue that a chatbot offers easy access to information for all types of donors and can support the transition of non-, first-time or lapsed donors to regular donors. By applying the design science research methodology, we have developed a chatbot for all donor types in South Africa and Ghana. We performed two design cycles, collaborating with experts from three blood services and grounding our research on existing and derived behavioural change models. The chatbot was positively evaluated in two workshops that included focus group discussions and online surveys.

Keywords: Chatbot, design science, blood donor management, behaviour change

Introduction

Blood donor management is an important part of blood logistics worldwide. As blood products cannot be produced artificially, but are crucial for many surgeries and treatments, sufficient and regular voluntary blood donations are necessary. In addition, blood products only have a short shelf-life, increasing the need for sufficient donations and an efficient use of resources. Fulfilling the demand for blood products on average days is already challenging in Ghana and South Africa (Vermeulen et al. 2019) and COVID-19 additionally worsened the situation. According to the Western Cape Blood Service, in South Africa's Western Cape region alone, approx. 700 units of whole blood must be collected every day to meet the demand. Since COVID-19, on average only 500 units have been collected daily. In South Africa and Ghana, less than 1 % of the population donates blood nowadays (SANBS 2023). According to the National Blood Service Ghana, in 2021, approx. 70 % of blood donations in Ghana were made by first-time donors. In the study of Asamoah-Akuoko et al. (2021) with first-time donors in Ghana in 2015, only 3.1 % of first-time donors donated again within six months. 68 % stated that this was due to missing information on when and where to donate again or because the blood collection teams had not returned. Therefore, it is important for the Ghanaian blood service to either convince them to become regular donors and provide necessary information and access or make sure to recruit sufficient new donors every year to cope with lapsed donors. The same is true for the South African blood services. While according to the Western Cape Blood Service, the percentage of first-time donors is significantly lower in the Western Cape region in South Africa, still ca. 50 % of them will not come back. Before COVID-19, according to the blood service, the recruiters managed to get 8 % of blood donors from educational programs at school in 2019, which decreased to only 1 % in 2020. Digital tools like chatbots can support educational programs when on-site visits are not possible or it is less easy to inform and recruit potential donors about the importance and necessity of blood donations.

In summary, it is crucial for blood services in Ghana and South Africa to receive sufficient blood donations on a regular basis. It is important for them to recruit a significant number of new (young) donors every year to cope with donors becoming lapsed donors or an increase or change in demand. In addition, they must motivate more first-time and lapsed donors to become regular donors (again). Therefore, blood services need to find efficient ways how to reach, interact with and provide information to all (potential) blood donors.

Diederich et al. (2019) have shown that inducing behaviour can be achieved by conversations with a chatbot when designed appropriately. In addition, chatbots are already used in various domains such as customer service and enterprise applications in many industries like banking, air travel and entertainment, but also healthcare, to support users in searching for relevant information and to automate easy tasks like checking the calendar for booking an appointment (Morana et al. 2017). We therefore argue that chatbots present a promising tool for blood services as well. In South Africa, 93 % of the population has access to a smart- or mobile phone (Kyle and Silver 2019). Most of the South Africans have an internet connection and use messaging apps like WhatsApp (GeoPoll 2020). Chatbots integrated into websites or WhatsApp can therefore offer (potential) donors in South Africa and Ghana fast and easy access to information around the blood donation process, potentially reaching a vast majority of the population. Previous research started the design process of a chatbot for blood donors in Germany together with the potential benefits of offering such a chatbot to all types of donors (Müller and Reuter-Oppermann 2022b). Complete design guidelines for blood donation chatbots with a specific focus on South Africa and Ghana to promote behaviour change and motivate donations are still missing. Consequently, by applying the design science research (DSR) methodology, we want to address the following research question with our work:

How to design chatbots for potential blood donors to promote sustainable prosocial behaviour?

In this paper, we focus on a chatbot that can be integrated into a website, for example, supporting behaviour change in a mainly passive way. We additionally discuss how to extend the chatbot to act proactively, for instance when it is integrated into a blood donation app or accessible via the WhatsApp messaging service. The implementation of such a chatbot will be part of our future work.

The remainder of this paper is structured as follows. In the following section, we present the related work on blood donor behaviour and on existing blood donation chatbots. Then, the design science research project is summarised, followed by the description of the design requirements, design principles and the instantiation of the chatbot. The performed evaluations follow in the next section. The paper closes with a summary and an outlook on future work.

Foundations and Related Work

Models Predicting Potential Blood Donors' Behaviour

For the design of socio-technical systems for behavioural change, psychological insights are indispensable. In order to predict future behaviour of potential blood donors, the consideration of past behaviour as the one main determinant is key (Bagozzi 1981; Charng et al. 1988; Ferguson and Bibby 2002). This determinant allows a differentiated perspective of the individual donor groups, whose generally valid predictor is intention: the more experience in donating blood, the smaller the role intention plays in predicting behaviour (Bagozzi 1981; Bagozzi 1982; Charng et al. 1988; Ferguson and Bibby 2002; Giles and Cairns 1995; Godin et al. 2007). While regular donors made blood donation a habit, the prosocial behaviour of the other groups of donors depends mainly on how motivated they are to donate blood. According to the theory of planned behaviour (TPB), intention as a predictor of future behaviour is determined by attitude, subjective norm and perceived behavioural control (Ajzen 1991). In the context of blood donation, attitude reflects the individual's weighing of the advantages (such as saving lives and being a role model for others) and disadvantages (such as physical reactions and revealed ineligibility) of donating blood (Burditt et al. 2009). Subjective norm refers to the individual's perception of the degree of blood donation approval by significant others like family members and friends and perceived behavioural control to its judgement of how easy or difficult donating blood might be. However, due to the fact that other important determinants like past behaviour are not covered by the TPB, many extended versions of the theory exist (e.g., Godin et al. 2005). Despite these extensions, it is almost impossible to predict how an individual develops and progresses in changing its behaviour (Masser et al. 2008).

Another theoretical framework that allows this prediction is the transtheoretical model (TTM), which is also applicable to blood donor behaviour besides the commonly used TPB (Amoyal et al. 2013; Burditt et al. 2009; Ferguson and Chandler 2005). The main focus of the TTM is on the stages (SOC) and processes of change (POC) (Prochaska and DiClemente 1982). In the context of blood donation, the stage of change describes the willingness of an individual to give blood and the POC characterise facilitating transitions between these SOC. There are five SOC: precontemplation (not planning to donate blood, e.g., due to unawareness, lack of knowledge or resistance regarding blood donation), contemplation (thinking about, but not committing oneself to donate blood within half a year), preparation (after initiating first steps, ready to donate blood within next month), action (actively donating blood for half a year) and maintenance (loyally donating blood for at least half a year). In addition, there are ten POC, which can be divided into experiential and behavioural transition strategies: consciousness raising (e.g., recalling given information on blood donation), dramatic relief (e.g., being emotionally touched by the opportunity to save someone's life through own blood donation), environmental re-evaluation (assumption that all of the people in the world would live a better life if every individual, who is eligible, donated blood), self-re-evaluation (e.g., feeling anticipated regret if no blood donation is made despite being eligible) and social liberation (e.g., realising that apart from hospitals there are other places to donate blood). The latter involve counter conditioning (e.g., being distracted while donating blood), helping relationships (e.g., donating blood together with a friend), reinforcement management (e.g., having the feeling of a rewarding experience through blood donation), self-liberation (e.g., committing oneself to donate blood) and stimulus control (e.g., distributing blood donation stickers at home). According to Ferguson and Chandler (2005), the developmental process of a blood donor career starts with the experiential and ends with the behavioural strategies. The authors also demonstrated that blood donor career and SOC are consistent with each other. Non-donors (nd) are in the precontemplation stage, first-time (fd) and lapsed donors (ld) are in the contemplation/preparation stages and regular donors (rd) are in the action/maintenance stages (Ferguson and Chandler 2005).

To fully understand blood donor behaviour, it is important to consider both models (TPB and TTM), as research has shown that the determinants of the TPB relate to the SOC of the TTM, i.e., attitude has a higher relevance, for example, in precontemplation stage, whereas subjective norm and perceived behavioural control are more relevant in later stages (e.g., Vries et al. 1998). Regarding attitude towards blood donation, being in the precontemplation stage, the disadvantages of donating blood clearly outweigh the advantages, whereas in the preparation and action/maintenance stages the exact opposite occurs (Burditt et al. 2009). According to Burditt et al. (2009), between contemplation and preparation stages there is no clear trend in one direction or the other. While the disadvantages become less relevant across the SOC, the advantages become significantly more relevant at least for the first three SOC, as being in the action/maintenance stages probably means blood donation has already become a habit and consideration of the advantages is not important anymore (Burditt et al. 2009).

Regarding the tailored design of chatbots, in order to better understand the blood donor behaviour of our user group, the development of an adapted model in the form of user archetypes in addition to and on the basis of the interplay between the TPB and TTM is essential due to the inclusion of further intuitive aspects (Floyd et al. 2008). In contrast to the high abstractness of theories, these aspects, which are consistent with the literature (e.g., Bednall et al. 2013; Ferguson 1996; Muthivhi et al. 2015) and validated by our African blood service collaborators, are more real-world oriented. With this, each donor group is characterised by different motives, challenges and needs regarding blood donation that have to be addressed in the chatbot design to reach all types of donors (Table 1). When designed appropriately, Diederich et al. (2019) have shown that inducing behaviour can be achieved by conversations with a chatbot.

Blood Donation Chatbots

Since the first rule-based psychotherapeutic chatbot ELIZA (Weizenbaum 1966) was developed in the 1960s, interest in chatbots has been increasing permanently. Technological advances as well as the ease of using chatbots enabling interaction through text messages in natural language are two of the main reasons for chatbot popularity. Even though a software is in the back, they mimic human conversation (Dale 2016). Matching their scope of application, chatbots can be used domain-specifically (e.g., in customer service) or as a general-purpose technology like ChatGPT (Gnewuch et al. 2017). They support users in searching for relevant information and take over simple tasks like booking an appointment through their easy access

User Archetype	Stage(s)	Motives	Challenges	Needs
Non-Donor	Precontemplation	- Never thought about donating blood - No time - No interest - Inconvenience - Rumours and misconceptions - Religiosity - (Imagined) ineligibility	- General lack of knowledge/awareness - Lack of knowledge, e.g., need for blood, donor site - Fear a) Needles b) Sight of blood c) Physical reactions d) Contagion e) Discovering illness	- Education - Expectation and experience exchange - Planning support - Motivation - Encouragement
First-time Donor	Contemplation/ Preparation	- Perceived need for donation after catastrophic events - Curiosity - Non-financial incentives, e.g., health check, gift item	- Deferral (temporary) - Inconvenience - Negative service - Doubts about eligibility to donate again (timing) - Fear a) Needles b) Sight of blood c) Physical reactions	- Education - Expectation and experience exchange - Planning support - Motivation - Encouragement
Lapsed Donor	Contemplation/ Preparation	- Deferral (temporary or permanently) - (Imagined) ineligible health conditions - Inconvenience - Negative service experience - Physical reactions after donations	- Deferral (temporary) - Inconvenience - Negative service experience - Fear of physical reactions	- Education - Expectation and experience exchange - Planning support - Motivation - Encouragement
Regular Donor	Action/ Maintenance	- Altruism - Personal moral norms - Positive self-esteem - Recognition - Convenience - Habit of donating blood - Non-financial incentives, e.g., health check, gift item	- Inconvenience - Negative service experience - Physical reactions	- Reinforcement - Organisational streamlining

Table 1. User Archetypes for Persona-Based Design

to available systems (Morana et al. 2017). By running cost efficiently and offering short resolution times 24/7, chatbots offer advantages to their providers as well as their users (Gnewuch et al. 2017). Being easily implemented on websites and messenger platforms, often used as massaging apps on smartphones, for many different application cases, chatbots are spreading rapidly.

One of the main opportunities of using chatbots is that they are perceived as anthropomorphic, giving the feeling of a human contact (Verhagen et al. 2014). Following the "Computers Are Social Actors" paradigm (Nass et al. 1994), this is achieved through the incorporation of social cues as design features derived from interpersonal communication (e.g., small talk and emojis) (Feine et al. 2019; Gnewuch et al. 2017). Social cues trigger subconscious social responses by the users that positively or negatively influence the degree of interaction with a chatbot, depending on the user perceptions and expectations towards it (Nass and Moon 2000; Nass et al. 1994) and numerous different design elements to stimulate perceptions and expectations of anthropomorphism towards the chatbot exist. Unfortunately, no precise guidelines for the human-like design of chatbots exist to date, which makes it difficult to select appropriate design features (Feine et al. 2019). Nevertheless, for the systematic selection of social cues, Feine et al. (2019) developed a taxonomy that categorises human-like cues of conversational agents into verbal, visual, auditory and invisible social cues with ten subcategories in total.

We argue that well-designed chatbots might serve as a persuasive and natural way to support and educate donors and promote sustained blood donation behaviour. However, to the best of our knowledge, only few publications have addressed the topic of designing chatbots for blood donors and so far, no chatbot was developed for the use in South Africa or Ghana. Initiated by a Brazilian blood donation centre and motivated by a large number of redundant enquiries about the blood donation process, Roman et al. (2020) propose a chatbot represented in the blood donation centre's mascot that is accessible via Google Assistant to educate and mobilise blood donors. The aim of the chatbot is to provide a simpler and faster access to information and an easy communication exchange in an entertaining way compared to websites by using multimedia content like images or links to videos. The authors decided to use informal language and equipped the chatbot with the ability to suggest topics in order to continue the conversation and to support the user with choosing the next one. In our paper, we build upon their work by extending the evaluation by a second standard questionnaire specifically aimed at chatbots as well as the collection of detailed feedback by asking specific questions about our design. Even though Roman et al. (2020) made use of a well elaborated standard questionnaire to evaluate their designed chatbot, the extension is necessary, since this questionnaire is applicable to any tool and not devised for the specific evaluation of chatbots. Beyond this, with a coverage of only 30 intents their chatbot is rather limited (vs. in our case: 82 intents). In the evaluation, with six minutes, the time the users were able to test the chatbot was very short (vs. in our case: 30 minutes). In addition, the authors did not perform a focus group discussion to collect detailed feedback. Looking into practice, the start-up BloodLink built a chatbot prototype for blood donors in India that was accessible via the Facebook Messenger (Mayroth 2017). Users could ask the chatbot specific questions about blood or the blood donation process as well as schedule blood donation appointments in their immediate vicinity. If necessary, the chatbot was able to forward the conversation to an experienced blood donor or doctor. The chatbot of the Canadian Blood Services is also accessible via Facebook Messenger (Canadian Blood Services 2017). It also informs potential donors about blood donation and its benefits and motivates them to make blood donation appointments. The chatbot of the American National Red Cross named ”Clara”, which appears in the form of a female doctor, is integrated into the website of the blood service (The American National Red Cross 2023). It opens the conversation with a list of three options a user can choose from. Thus, it can either answer questions about eligibility and the blood donation process or make appointments. It enables free text input but also provides guidance to the user in the form of buttons and website links. If the chatbot does not know the answer to the user’s question, it can provide the user with the contact details of the blood service for further clarification. Even though the aforementioned chatbots are partly represented by avatars and names, overall they only have very few social cues (e.g., no use of small talk, response delays and emojis). Additionally, they are all able but limited to answer specific questions about blood donation and its process as well as making appointments. Therefore, in contrast to them, we not only equip the chatbot with a broader set of social cues, but we are the first who plan its integration into a blood donation app to provide the user with a holistic set of features in order to ensure long-term donor engagement. Besides chatbot push notifications and reminders, these features include, for example, the donor questionnaire incorporated digitally as well as gamification elements and performance records, as listed in our previous work about blood donation app design (Müller and Reuter-Oppermann 2022a). Considering the cultural perspective and comparing the perceived persuasiveness of South Africans having a more individualist blood donation culture with the one of Ghanaians having a more collectivist culture, we found that for the former, personal features such as reminders or rewards are more important, while for the latter, social features such as normative influence and social role dominate (Müller and Reuter-Oppermann 2023).

Overall, with respect to the literature and existing chatbots, it can be stated that thorough evaluations and design aspects to ensure a sustainable positive influence on the user’s behaviour (i.e., transition to and retention of regular blood donors) are missing. In order to address this lack of design knowledge on blood donation chatbots, we argue that it is suitable to apply the DSR methodology.

Design Science Research Project

Design science research is used to design and evaluate new and innovative artefacts, e.g., models or software products, in iterative cycles in order to solve practical problems (Hevner 2007). According to Hevner (2007), besides addressing practical relevance, the effective incorporation of scientific rigor is equally important to contribute to both, technological innovations and information systems research knowledge. We argue that this research approach is particularly suited to answer our research question, since we address the real-world problem of matching demand and supply of blood donations by designing and evaluating a chatbot in the specific context of potential blood donors in Ghana and South Africa to increase their willingness to give blood. Increasing the relevance of our research, we collaborate with experts from three blood services from those regions, the National Blood Service Ghana (NBSG), the Western Cape Blood Service (WCBS) in South Africa and the South African National Blood Service (SANBS) that provides its service to all of South Africa’s provinces except of the Western Cape province. Accordingly, all three serve for providing additional requirements and for evaluating our resulting design as they have years of experience with and in-depth knowledge of their (potential) blood donors. Increasing the rigor of our research, we include behavioural

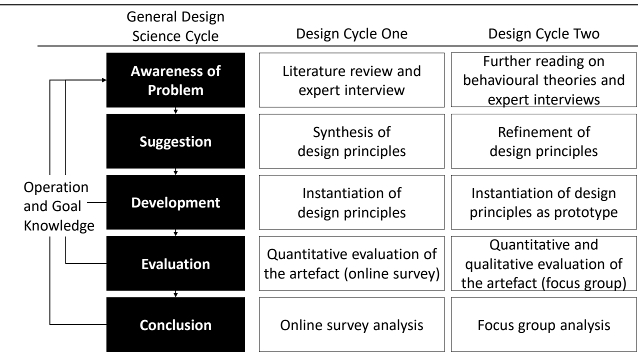

Figure 1. Design Cycles with Respective Research Activities

change models for grounding our design as suggested by Hevner (2007) to make use of a complementary research cycle by bridging claims to design science as well as behavioural science. As shown in Figure 1, the DSR framework proposed by Kuechler and Vaishnavi (2008) builds the basis of our research project with two subsequent design cycles.

In our first design cycle as part of our previous work, we derived three design principles (DPs) for chatbots for potential blood donors increasing their willingness to donate blood grounded on requirements extracted from extant literature on chatbots, blood donor mobilisation and management solutions and blood donor behaviour (Müller and Reuter-Oppermann 2022b). In addition, we did an interview with an expert of the blood donor behaviour of African minorities in Western Europe as we did not have access to our project partners at that point due to COVID-19 and asked about (1) the general opinion on chatbots, (2) challenges regarding communication and information exchange between blood services and donors and (3) application scenarios and usefulness of chatbots. With Botframe, a platform for creating previews of chatbot interactions, we subsequently instantiated our three DPs in a chatbot prototype. We then evaluated parts of this prototype embodying some of our identified design requirements (DRs) together with ten fictional blood donation scenarios representing the rest of our requirements (e.g., blood donation appointment booking) in an end user online survey. The survey revealed that with regard to the individual donor types (i.e., nd, fd, ld, rd), to embrace the "one size fits all" principle, all of our identified requirements have to be considered for the development of the chatbot, because each donor group has different preferences on what and how exactly a chatbot should provide support to them. A detailed description of our first design cycle including our research results is presented in (Müller and Reuter-Oppermann 2022b).

We started our second design cycle with further reading on behavioural theories and two semi-structured interviews conducted in April 2022 with our partners from South Africa on site. One interview was performed at WCBS in Cape Town with the Head of Marketing and PR and the other one at SANBS in Johannesburg with the Head of Marketing and PR and with the Senior Manager Business Intelligence, who responded to the respective questions. In both interviews we again asked about aspects (1) - (3) with an emphasis on the potential blood donors' needs extracted from our derived behavioural change model regarding the four different user archetypes. We conducted the guided interviews in order to refine as well as identify design requirements in addition to those defined in the first design cycle. We added one design requirement (DR3.6) to address the issue of stage-matched interventions described in the interviews. For example, one interviewee stated that the blood service usually addresses lapsed donors by giving an explanation why the particular person is needed and that s/he can reply via SMS in case s/he wants to go back into the donor pool. Our interviewee believed this to be a cumbersome process which can be substituted by a chatbot facilitating direct follow-up conversations to foster donor engagement. After the refinement of the third design principle, we instantiated our three DPs in a chatbot prototype developed with Google Dialogflow providing natural language processing (NLP) capabilities for user intent detection and a custom-built web interface providing convenient access. We then evaluated our prototype with two workshop groups interacting with our chatbot, quantitatively by rating the chatbot with the help of standardised questionnaires (i.e., the User

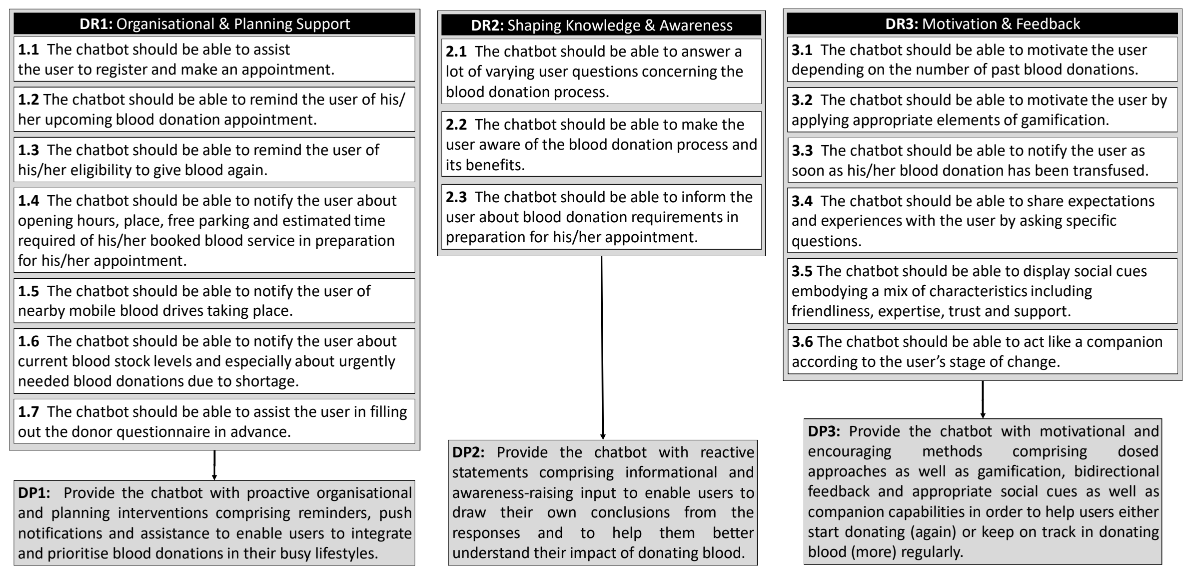

Figure 2. Design Requirements and Design Principles for Blood Donation Chatbots (based on Müller and Reuter-Oppermann (2022b))

Experience Questionnaire and the Chatbot Usability Questionnaire) and qualitatively by conducting focus group discussions. The first workshop in September 2022 took place in Ghana with 19 representatives of the NBSG and the second one in October 2022 in Berlin with nine representatives of all of our African partners. Each workshop served for the assessment of potential blood donors' perceptions and attitudes towards the chatbot as well as getting feedback for its further improvement.

Evaluating the chatbot with experts from the blood services was a necessary and important step in the design process. Due to the importance of donor education provided by the blood services, we aimed to evaluate the usefulness of the chatbot in supporting them in this task. In addition, it was a prerequisite in order to gain access to donors for a future evaluation in a subsequent design cycle.

Designing Chatbots for Blood Donor Behaviour Change

In order to induce changes in blood donor behaviour, it is essential to know barriers and motivators of potential blood donors grounded on the behavioural change models presented in the foundations. With this, we ensure a user-centred design of the chatbot following the design principles shown in Figure 2. An instantiation of the design principles can be found in Figure 3.

As stated by the blood services, donor education is a challenging issue and therefore an important task for a chatbot. The online survey in our first design cycle revealed that, when comparing all donor groups, first-time donors would benefit the most from an additional source of information providing easy and fast access, probably due to their uncertainty before and after donating blood for the first time (**DR2.1**) (Müller and Reuter-Oppermann 2022b). Often many questions like "What happens with my blood?" and "What are reasons for being temporary deferred and what does that exactly mean?" arise, for which clarification is needed. With the help of an easily accessible tool to quickly answer these questions, return rates of first-time donors could be much higher (Piliavin and Callero 1991). For example, according to WCBS, of the approx. 30 000 donors recruited on average each year, only about half of them return to donate a second time. In Ghana, according to Asamoah-Akuoko et al. (2021), the rate for first-time donors coming back within six months is even lower (15.2 %), presumably mainly due to the fact that ca. 70 % of the blood donors there are family replacement donors (i.e., donations are made to cover the blood products needed by a family member or friend), a concept not used in South Africa. Additionally, the Head of Marketing and PR of WCBS

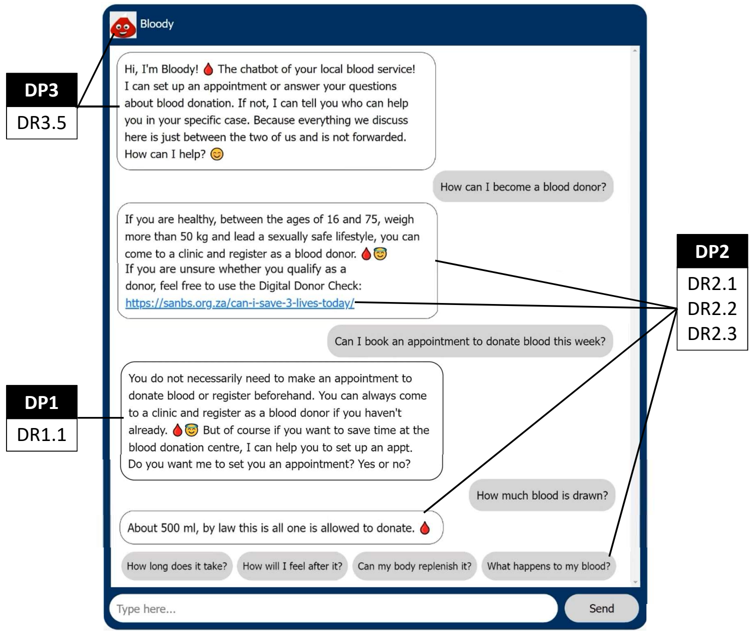

Figure 3. Instantiated Design Principles

reported that ca. 90 % of the questions reaching customer service every day are standard questions. In her opinion, for an easier and faster communication and information exchange, the chatbot should be able to answer simple questions like “Can I donate blood?” and “Where can I donate blood?”. To address this, we consulted the FAQ websites of WCBS, SANBS and NBSG (extended by some questions of the known FAQ websites of the German Red Cross to guarantee completeness considering basic questions) and modelled the different extracted user intents for approaching the chatbot on Google Dialogflow (i.e., 82 intents in total). In order to ensure a standardised conversational style, we reviewed the respective training phrases (i.e., user requests) and responses of the chatbot. However, in the interview with SANBS, the Senior Manager Business Intelligence indicated that even if the FAQ are covered by the chatbot, medical questions are still at concern (e.g., when the chatbot motivates to donate blood despite taking a particular medication, but at the blood service it is prohibited, frustrating deferrals might be the consequence SANBS wants to avoid). The Head of Marketing and PR of WCBS also highlighted to put emphasis on medical issues (e.g., clarification of Hb level to avoid people coming in vain) and suggested that in case of specific user requests, the chatbot should be able to point to the right contact person instead. To address this, we generalised the chatbot responses for this kind of requests and added telephone numbers as well as e-mail addresses of the blood service, carefully selected depending on the specific case. For both blood services, another important issue to consider is the unpredictable course of the COVID-19 pandemic. The online survey in the first design cycle revealed that, comparing all donor groups, mostly regular donors tend to consult a chatbot to answer questions about blood donation in times of COVID-19 (Müller and Reuter-Oppermann 2022b). In this case, they probably benefit most from a chatbot, because due to their frequent donation centre visits, they are most affected by constantly changing regulations (Müller and Reuter-Oppermann 2022b). Hence, we additionally integrated user intents extracted from the digital COVID-19 FAQ booklet of SANBS on Google Dialogflow (i.e., eight out of 82 intents refer to COVID-19). For addressing the high variability of user input, we also created a fallback response in case intent detection was not possible.

The blood service employees interviewed by Batis and Albarrak (2021) in the context of a blood donation app suggested to include additional educational materials for awareness-raising and increasing the attractiveness of donating blood (**DR2.2**). Roman et al. (2020), for example, have demonstrated that this is also possible within a blood donation chatbot by including multimedia content like images and links. Therefore, depending on the questions, we added images and website links of SANBS to the responses of the chatbot for extra information that is optionally accessible for the users, while the chatbot itself already provides as many details as possible. For our prototype, we decided to include links forwarding to the website of SANBS, because compared to WCBS and NBSG it is the blood service with the most donors and the best laid out website. The former is also the reason why we tailored the chatbot responses to the South African context in this work, because depending on the country, the requirements of the blood services differ to some extent (e.g., regarding the age range for donating blood). In future work, additional tailored artefacts for WCBS and NBSG will be built. In relation to the Senior Manager Business Intelligence's statement of the aim to avoid frustrating deferrals due to potential losses of blood donors, the blood service employees in the study of Batis and Albarrak (2021) also emphasised the importance of providing guidance on pre-donation requirements to prevent coming in vain (**DR2.3**). Hence, we not only integrated website links to raise awareness but also to check the own eligibility to donate blood in the form of a self-test by clicking through the individual blood donation criteria and deciding whether they are met or not (e.g., "Do you weigh 50 kg or more?"). Overall, the integration of this kind of multimedia content might be particularly useful to non-donors as well as first-time donors.

The online survey of our first design cycle revealed that comparing all donor groups, mostly lapsed donors tend to use a chatbot to set up an appointment (**DR1.1**), probably because they need more planning support (Müller and Reuter-Oppermann 2022b). As the Head of Marketing and PR of WCBS suggested to include appointment booking for plasma donations as a feature to the chatbot, which is predominantly done by regular donors, they might also benefit from making appointments quickly and easily. Thus, for enabling appointment booking, we incorporated entities on Google Dialogflow, which are annotations for parameters like time and place queried during the chatbot conversation as well as contexts allowing sequential dialogues in case user input just represents a simple "yes" or "no".

Considering the way the chatbot should provide communication and information exchange with regard to the individual donor groups, the Head of Marketing and PR of WCBS proposed to generally enable pleasant human-like interactions (**DR3.5**). Based on the taxonomy of Feine et al. (2019) mentioned in the foundations, we purposefully selected a set of social cues for the human-like design of our text-based conversational agent (i.e., elimination of auditory category as well as all subcategories belonging to embodied conversational agents). Concerning verbal cues, referring to the content, Amershi et al. (2019) as well as Weber and Ludwig (2020) for instance recommend that right at the beginning of an interaction a chatbot should be able to clarify what it is capable of and what not, because without initial clarification, users usually do not intuitively know their purpose when interacting with them for the first time. Additionally, in terms of trust and privacy issues, the chatbot should be able to explain the handling of user information (Weber and Ludwig 2020). Referring to the style, we for example decided to apply informal language to appeal especially to younger people who, with regard to blood donation, are underrepresented in South Africa, where more than half of all blood donors are over 30 years old (Vermeulen et al. 2019). Considering visual cues, Roman et al. (2020) for instance finds empirical evidence in an experiment that the blood donation chatbot representation in the form of an avatar (in this case mascot of the blood donation centre in Brazil) increases user satisfaction as well as trust, while it decreases users' inhibition threshold. With regard to invisible social cues, Gnewuch et al. (2017) for example demonstrated that dynamic response times also contribute positively to the chatbot's anthropomorphic perception despite longer latency times for the user by using response delays for simulation of the chatbot's thinking and typing.

To address the new design requirement **DR3.6** that the chatbot should be able to act like a companion according to the user's stage of change, we plan to integrate the chatbot into a blood donation app allowing additional functionalities for the individual types of donors (Müller and Reuter-Oppermann 2022a). Müller and Reuter-Oppermann (2022a) substantiated the functional compatibility of a blood donation app and a chatbot and explained how their interaction positively influences blood donor behaviour change. As registration within the app is directly followed by a short self-test comprising four blood donation questions to determine the user's initial stage of change, the chatbot can behave and react in a way that is fully adapted to the user. Like a companion, the chatbot can provide stage-matched guidance and feedback to the user to promote the developmental process of the user's blood donor career (Amoyal et al. 2013). Following this argumentation, regarding the other proactive chatbot interventions (i.e., **DR1.2 – DR1.7** and **DR3.1 –**

DR3.4), we also plan to integrate them into a blood donation app in future work.

Evaluation

Evaluation Methodology

According to Maedche et al. (2019), a chatbot is a socio-technical information system consisting of a trilogy of user, task and technology. However, we do not focus on the technical but on the user-centred design in this work. This means not only to fulfil the chatbot's purpose of being able to mobilise and manage potential blood donors, but also to elicit a positive user experience (UX) towards the chatbot. The latter is essential for being sustainably accepted, to catch on and achieve pervasiveness in order to mobilise as many potential donors as possible. According to Bhattacherjee and Lin (2015), three constructs directly and positively influence IS continuance: subjective norm, perceived usefulness and satisfaction. User experience encompasses the latter two, including many other quality attributes besides usability (e.g., functionality) (Laugwitz et al. 2008). Consequently, we evaluated our chatbot by conducting a UX assessment via two exploratory focus group workshops with our African partners to get feedback on how the chatbot is perceived and how it can be improved (Tremblay et al. 2010), responding also to the need of more real-world evaluations of DSR artefacts (Peffers et al. 2012; Venable et al. 2016). The feedback of the experts is important for the design project at this stage, as they have profound knowledge of and experience with blood donor's needs and questions and can also comment on the correctness of the answers and information provided by the chatbot.

Based on the UX framework of Zarour and Alharbi (2017), similar to Roman et al. (2020), we used the User Experience Questionnaire (UEQ), which in contrast to other questionnaires equally considers pragmatic as well as hedonic product quality aspects by using the UX framework of Hassenzahl (2001) (Laugwitz et al. 2008). Since this questionnaire serves to assess the general UX of systems for human-computer interaction, we additionally applied the Chatbot Usability Questionnaire (CUQ), which is based on the chatbot UX categories elaborated by Martín et al. (2017) and specifically tailored to the UX assessment of chatbots (Holmes et al. 2019). The UEQ is grounded on six scales and 26 items representing these scales: attractiveness (i.e., extent of likeability of the product), perspicuity (i.e., extent of familiarity with the product), efficiency (i.e., extent of usefulness of the product), dependability (i.e., extent of user control), stimulation (i.e., extent of motivation to use the product), and novelty (i.e., extent of innovativeness of the product). They are rated through item pairs of antonymous adjectives describing the product (e.g., "annoying" vs. "enjoyable"). Seven stage scales are provided for each pair allowing the user to respectively select its level of agreement with one of the presented items. The CUQ is composed of 16 statements, of which half reflect positive aspects and the other half negative aspects of chatbot usability. For each statement, the respondents decide on their level of agreement via a five-point Likert scale (i.e., from 1 = "strongly disagree" to 5 = "strongly agree"). In order to get qualitative feedback on our proposed design, we also conducted focus group discussions with our African collaborators.

In our first workshop in Ghana, the focus group consisted of employees of one specific blood service (i.e., NBSG). In the second, the focus group was fairly mixed with two participants from NBSG, two from SANBS, three from WCBS and two additional external experts. The former served for the consideration of experts from all the different departments (e.g., IT, laboratory and donor care), potential donors might have related questions about (e.g., "What is my blood tested for?"). Out of the 19 representatives of the NBSG, we also considered a fairly balanced gender (i.e., m = 8 (42.1 %) and f = 11 (57.9 %)) as well as age mix (from 29 to 58 years). Overall, with this purposeful selection, we aimed to ensure comprehensive feedback from diverse perspectives and to investigate the differences regarding demographic variables.

Each workshop lasted one and a half hours in total and was structured as follows. First, each participant had to read the instructions comprising a brief definition of chatbots, our objectives and the procedure of the workshop. After the participants' agreement to our privacy policy as well as the use of Google's Dialogflow (for interaction with our chatbot) and Forms (to conduct our online survey) services, they had to answer three control questions to make sure that they understood our chatbot experiment correctly. Then, the interaction with the chatbot started. For 30 minutes, each participant using their own laptop or smartphone freely interacted with our chatbot empathising with the role of being a potential blood donor with a lot of different questions in mind. After the interaction time was over, the chatbot provided a survey link the participants had to open to answer questions related to their demographics, the UEQ, the CUQ and their general feedback. In the remaining 45 minutes, we conducted the focus group discussion consisting of general questions (e.g., "What did you like or dislike about the chatbot?"), more specific questions concerning our three DPs (e.g., "Do you think that the chatbot could motivate potential blood donors to donate blood?"), questions related to the output of the chatbot (e.g., "What did you think of the answers the chatbot gave to your questions?") as well as questions specifically considering our human-like design (e.g., "What did you think of the name and look of the chatbot? What would you change and how?"). With the consent of the participants, the focus group discussion was recorded and transcribed afterwards for further analysis. In the next section, we present the results of our two exploratory focus group workshops and discuss the feedback in detail.

Results and Discussion

Overall, the blood donation experts who participated in our focus group workshops rated the UX of our chatbot rather positively. In general, we revealed this by conducting data analyses of our two questionnaires (i.e., UEQ and CUQ), for which analytical tools are provided to obtain scores for comparison. Considering the UEQ, we also compared the association between scores and demographic variables by performing the Mann-Whitney-U-Test with the use of IBM SPSS Statistics 29.0.

Regarding the seven-stage semantic differential scales of the UEQ, the gradations are equivalent to -3, -2, -1, 0, 1, 2, 3 with scores ranging from -3.000 to 3.000 according to the user ratings. If the average rating is lower than –0.800, the UX is perceived negatively, whereas it is perceived positively when the mean is higher than 0.800 and neutral if in between.

Table 2 shows the results per item pair and the overall scales. Considering the results of the Ghana workshop, all of the six scales as well as item pairs were rated with a mean higher than 0.800 besides the pair with the items "unpredictable" and "predictable" (-0.105). Here, the user responses are distributed fairly with a slight tendency towards the chatbot being perceived as unpredictable. One possible reason behind that might be that a chatbot represents a "blank canvas where the content and features of the underlying service are mostly hidden from the user" (Følstad and Brandtzæg 2017, p. 41). The unfamiliarity with chatbot features is especially true for users who have never used a chatbot before. This is why the distribution of the user responses of the workshop conducted in Berlin who indicated that they used a chatbot at least once, is different compared to the one in Ghana, where 11 out of 19 participants stated that they have no experience in chatbot usage. Accordingly, the mean of the dependability scale (1.289) is the lowest compared to the other overall scales of the Ghana workshop (Table 2). The comparatively low standard deviations of the other item pair ratings support the positive UX assessment of the chatbot. Overall, the attractiveness as well as the stimulation scale achieved the highest means (each 2.211). The rating of the attractiveness scale was mainly influenced by the chatbot being perceived as good (2.632 = highest mean) and the stimulation scale by the chatbot being perceived as interesting (2.368 = second highest mean, same for being perceived as friendly belonging to the attractiveness scale). This lets us assume that the participants of the workshop in Ghana generally liked the chatbot and that the chatbot motivated them to interact with it, which overall indicates a user-friendly design. The distribution of the responses shows that the participants of our workshop in Berlin rated the chatbot slightly more negatively compared to them in Ghana, mainly for the item pairs representing the novelty scale (i.e., "conventional/inventive" (mean = 0.566), "usual/leading edge" (mean = 0.889) and "conservative/innovative" (mean = 1.111)). Therefore, it is obvious that the mean of the novelty scale (0.972) is the lowest compared to the other overall scales of the Berlin workshop. This might be due to the fact that all of our nine participants were already familiar with the usage of chatbots and it was nothing new to them, which was not the case in Ghana as mentioned before. On average, the item pair "conventional/inventive" was rated below 0.800 and the chatbot was perceived as being rather slow (0.111). Regarding the incorporation of dynamic response times for anthropomorphic design, this probably means that the participants of the Berlin workshop were more in favour with prompt answers, presumably because they are used to it from the usage of other "traditional" chatbots. The six scales and the remaining item pairs were rated with a mean above 0.800, indicating that the UX of the chatbot was generally perceived positively by the participants. Overall, the perspicuity scale achieved the highest mean (1.944), which was mainly influenced by the chatbot being perceived as easy to learn (2.444 = highest mean). However, the results of

	Ghana Workshop (N = 19)		Berlin Workshop (N = 9)		
Semantic Differential Rating	Mean ± SD	95% CI	Mean ± SD	95% CI	Scale/Quality
annoying/enjoyable	2.053 ± 0.780	0.351	2.000 ± 0.707	0.462	Attractiveness
not understandable/understandable	2.000 ± 0.943	0.424	1.889 ± 0.928	0.606	Perspicuity
dull/creative	2.263 ± 0.806	0.362	1.333 ± 1.225	0.800	Novelty
difficult to learn/easy to learn	2.316 ± 0.820	0.369	2.444 ± 0.882	0.576	Perspicuity
inferior/valuable	2.158 ± 0.834	0.375	1.556 ± 1.878	1.227	Stimulation
boring/exciting	2.158 ± 0.765	0.344	1.556 ± 1.014	0.662	Stimulation
not interesting/interesting	2.368 ± 0.831	0.374	2.000 ± 0.707	0.462	Stimulation
unpredictable/predictable	-0.105 ± 1.761	0.792	1.111 ± 1.167	0.762	Dependability
slow/fast	1.842 ± 0.834	0.375	0.111 ± 1.764	1.152	Efficiency
conventional/inventive	1.737 ± 0.991	0.446	0.556 ± 1.667	1.089	Novelty
obstructive/supportive	2.211 ± 0.713	0.321	1.889 ± 1.054	0.689	Dependability
bad/good	2.632 ± 0.496	0.223	1.444 ± 1.333	0.871	Attractiveness
complicated/easy	2.158 ± 1.214	0.546	2.333 ± 0.707	0.462	Perspicuity
unlikable/pleasing	2.000 ± 0.943	0.424	1.778 ± 0.667	0.436	Attractiveness
usual/leading edge	1.105 ± 1.487	0.669	0.889 ± 1.764	1.152	Novelty
unpleasant/pleasant	2.053 ± 0.705	0.317	1.667 ± 0.866	0.566	Attractiveness
not secure/secure	1.632 ± 1.499	0.674	1.000 ± 1.323	0.864	Dependability
demotivating/motivating	2.158 ± 1.015	0.456	1.222 ± 1.093	0.714	Stimulation
does not meet expectations/meets expectations	1.421 ± 1.017	0.457	1.222 ± 1.481	0.968	Dependability
inefficient/efficient	1.895 ± 0.937	0.421	1.556 ± 0.882	0.576	Efficiency
confusing/clear	2.053 ± 0.848	0.381	1.111 ± 1.167	0.762	Perspicuity
impractical/practical	2.053 ± 0.911	0.410	2.111 ± 0.601	0.393	Efficiency
cluttered/organised	2.053 ± 0.970	0.436	1.667 ± 0.866	0.566	Efficiency
unattractive/attractive	2.158 ± 1.119	0.503	1.889 ± 1.054	0.689	Attractiveness
unfriendly/friendly	2.368 ± 0.761	0.342	2.111 ± 0.928	0.606	Attractiveness
conservative/innovative	2.105 ± 0.809	0.364	1.111 ± 1.900	1.242	Novelty
Overall Attractiveness Scale	2.211 ± 0.608	0.273	1.815 ± 0.738	0.482	Attractiveness
Overall Perspicuity Scale	2.132 ± 0.704	0.317	1.944 ± 0.808	0.528	Pragmatic
Overall Efficiency Scale	1.961 ± 0.723	0.325	1.361 ± 0.830	0.542	Pragmatic
Overall Dependability Scale	1.289 ± 0.591	0.266	1.306 ± 1.006	0.657	Pragmatic
Overall Stimulation Scale	2.211 ± 0.652	0.293	1.583 ± 0.976	0.638	Hedonic
Overall Novelty Scale	1.803 ± 0.616	0.277	0.972 ± 1.411	0.922	Hedonic

SD = Standard deviation, CI = Confidence interval

Table 2. Results of the UEQ for the Two Workshops

the workshop conducted in Berlin should be treated with caution due to the high standard deviations of the item pair ratings and the rather low sample size. Table 3 summarises the statistical analysis of the UEQ scales. The results of the Berlin workshop clearly show that the chatbot was generally rated more positively by participants not belonging to the South African blood services (i.e., SANBS and WCBS). Compared to the latter, they perceived the chatbot as more attractive as well as innovative. Same is true for the younger participants. This was not the case for the South African partners, probably because they are mostly composed of IT specialist and probably very experienced when it comes to technology. Moreover, in contrast to the rest of our African partners, they might have been biased due to our former visits on site when we were already discussing about the chatbot. For workshop participants in Ghana, regardless of their age or chatbot usage experience, the UX of the chatbot was rated equivalently positive. Regarding the perception of the chatbot's usefulness (i.e., efficiency scale), differences in demographic variables play a role. This means the chatbot was perceived as more efficient by women as well as participants with a lower educational qualification and a lower approach to new technologies. Additionally, the latter also perceived the chatbot as more attractive compared to participants being technology-savvy, probably because the conversational user interface is easily accessible via natural language.

Regarding the 16 statements of the CUQ, the odd-numbered relate to the positive usability aspects of the chatbot, which are supported by the users when the level of agreement is rather high (i.e., towards 5 = "strongly agree"). The even-numbered statements relate to the negative usability aspects of the chatbot, which are not supported by the users when the level of agreement is rather low (i.e., towards 1 = "strongly disagree"). Comparing the level of agreements of the participants from the Ghana and Berlin workshop shown in Table 4, it can easily be seen that they both agreed to the positive usability aspects and rather disagreed to the negative ones with rating scores being quite similar. Overall, this means that the partici-

	Variable	n(%)	Attractiveness	Perspicuity	Efficiency	Dependability	Stimulation	Novelty
Ghana Workshop (N = 19)	Gender		p = 0.699	p = 0.200	p = **0.029***	p = 0.886	p = 0.886	p = 0.886
	male	8(42)	2.146 ± 0.511	1.875 ± 0.643	1.594 ± 0.820	1.094 ± 0.231	2.250 ± 0.607	1.938 ± 0.442
	female	11(58)	2.258 ± 0.302	2.318 ± 0.351	2.227 ± 0.181	1.432 ± 0.414	2.182 ± 0.339	1.705 ± 0.348
	Age		p = 0.485	p = 0.686	p = 0.114	p = 0.486	p = 0.200	p = 0.686
	>45	5(26)	2.333 ± 0.097	2.050 ± 0.044	1.800 ± 0.388	1.750 ± 0.688	2.300 ± 0.106	1.700 ± 0.075
	$\leq$ 45	14(74)	2.167 ± 0.474	2.161 ± 0.669	2.018 ± 0.591	1.125 ± 0.161	2.179 ± 0.552	1.839 ± 0.496
	Educational qualification		p = 0.132	p = 0.886	p = **0.029***	p = 1.000	p = 0.486	p = 0.886
	Master or higher	9(47)	2.019 ± 0.260	2.111 ± 0.502	1.833 ± 0.281	1.278 ± 0.148	2.083 ± 0.219	1.667 ± 0.172
	Bachelor or lower	10(53)	2.383 ± 0.439	2.150 ± 0.544	2.075 ± 0.765	1.300 ± 0.567	2.325 ± 0.626	1.925 ± 0.570
	Approach to new tech.		p = **0.015***	p = 0.200	p = **0.029***	p = 1.000	p = 0.114	p = 0.486
	tech.-savvy	8(42)	1.854 ± 0.464	2.000 ± 0.696	1.750 ± 0.625	1.375 ± 0.429	1.938 ± 0.638	1.531 ± 0.276
	not tech.-savvy	11(58)	2.470 ± 0.166	2.227 ± 0.381	2.114 ± 0.442	1.227 ± 0.318	2.409 ± 0.216	2.000 ± 0.388
	Chatbot used before		p = 0.394	p = 0.200	p = 0.114	p = 1.000	p = 0.057	p = 1.000
	yes	8(42)	2.104 ± 0.690	2.313 ± 0.728	1.875 ± 1.143	1.094 ± 0.249	2.063 ± 0.817	1.875 ± 0.661
	no	11(58)	2.288 ± 0.167	2.000 ± 0.338	2.023 ± 0.131	1.432 ± 0.401	2.318 ± 0.164	1.750 ± 0.213
Berlin Workshop (N = 9)	Gender		p = 0.937	p = 0.886	p = 0.343	p = 0.057	p = 0.057	p = **0.029***
	male	6(67)	1.806 ± 0.327	1.917 ± 0.517	1.125 ± 0.394	1.000 ± 0.775	1.333 ± 0.892	0.542 ± 2.110
	female	3(33)	1.833 ± 1.361	2.000 ± 1.313	1.833 ± 1.271	1.917 ± 1.271	2.083 ± 1.021	1.833 ± 1.021
	Age		p = **0.041***	p = 0.886	p = 0.200	p = 0.200	p = **0.029***	p = **0.029***
	>45	6(67)	1.694 ± 0.216	1.958 ± 0.535	1.042 ± 0.210	1.083 ± 0.917	1.292 ± 0.785	0.375 ± 1.594
	$\leq$ 45	3(33)	2.056 ± 1.509	1.917 ± 1.271	2.000 ± 1.313	1.750 ± 1.313	2.167 ± 1.083	2.167 ± 0.771
	Educational qualification		p = 0.180	p = 0.343	p = 0.486	p = **0.029***	p = 0.200	p = 0.057
	Master or higher	7(78)	1.881 ± 0.590	2.107 ± 0.476	1.464 ± 0.842	1.571 ± 0.557	1.750 ± 0.792	1.179 ± 2.348
	Bachelor or lower	2(22)	1.583 ± 0.681	1.375 ± 1.531	1.000 ± 0.125	0.375 ± 2.531	1.000 ± 2.000	0.250 ± 0.500
	Approach to new tech.		p = 0.180	p = 0.486	p = 0.686	p = 0.886	p = 0.686	p = 0.114
	tech.-savvy	7(78)	1.762 ± 0.378	2.000 ± 0.333	1.214 ± 0.384	1.357 ± 0.164	1.607 ± 0.518	0.857 ± 1.705
	not tech.-savvy	2(22)	2.000 ± 2.000	1.750 ± 3.125	1.875 ± 2.531	1.125 ± 7.031	1.500 ± 4.500	1.375 ± 5.281
	Country		p = **0.002***	p = 0.114	p = 0.200	p = 0.057	p = **0.029***	p = **0.029***
	South Africa	5(56)	1.367 ± 0.339	1.600 ± 0.831	0.950 ± 0.075	0.750 ± 0.781	0.950 ± 0.575	0.000 ± 0.938
	rest	4(44)	2.375 ± 0.248	2.375 ± 0.188	1.875 ± 1.104	2.000 ± 0.500	2.375 ± 0.271	2.188 ± 0.516
	Position		p = 0.699	p = 0.200	p = 0.486	p = 1.000	p = 0.200	p = **0.029***
	IT	3(33)	1.722 ± 0.176	2.250 ± 0.063	1.083 ± 0.083	1.250 ± 0.063	1.250 ± 0.563	-0.333 ± 0.896
	rest	6(67)	1.861 ± 0.794	1.792 ± 0.935	1.500 ± 1.000	1.333 ± 1.592	1.750 ± 1.200	1.625 ± 1.294

Table 3. Statistical Analysis of the UEQ Item Scales (Results Expressed as Mean +/- SD, *p $\leq$ 0.050)

pants of both workshops rated the usability of the chatbot rather positively than negatively. However, the participants of the Berlin workshop slightly tended to perceive the chatbot as robotic (no. 2) and they were undecided if the chatbot failed to recognise at least some of their inputs (no. 10).

With the help of the two focus group discussions, we gained further valuable insights form the participants on aspects that are already good and on those that could be improved in future research. They perceived the chatbot as very easy, interactive, supportive and user-friendly. They described the language of the chatbot as useful, simple and concise to the point. They are confident that the chatbot will particularly add value in educating donors, an issue that blood services have very much on target. They perceived the chatbot as very clear and helpful with the opportunity to get some further details. Even though the answers were not matched a hundred percent correctly to the users' intent, the participants appreciated that the users still learn something by interacting with the chatbot. As improvements, they asked for a broader variation in chatbot responses, which are the same for most of the questions when being asked several times. Additionally, when the same question is being asked in a different way, the participants reported that the chatbot sometimes had problems giving the same correct answer again. Therefore, the NLP capabilities need to be improved by providing more training phrases on Google Dialogflow with the help of the conversation logs from the workshops. The participants stated that for more complex questions, it was harder for the chatbot to provide the right answers. Additionally, the Head of Donor Management of WCBS suggested that the chatbot should be able to provide specific answers to certain deferral issues (e.g., surgery, medication, travelling abroad) by asking follow-up questions such as "What kind of medication are you taking?", which also contributes to a more interactive conversation. Moreover, they pointed out (again) that people of African culture have very strong beliefs and perceptions, including myths and misconceptions preventing them from donating blood, and that the chatbot could put more emphasis on this. According to our participants, this is particularly important for the chatbot to motivate potential blood donors to give blood. In their opinion, the chatbot needs to be more persuasive addressing the barriers that keep people from donating blood. They also highlighted that the chatbot could personalise messaging much more to a potential donor by asking for its name, for example, or thanking them for donating blood. Nevertheless, they stated that they would use the chatbot again and especially recommend it to first-time donors, as they will probably learn the most as well as to potential donors who fear that they will be deferred by the blood service. Regarding the

No.	Statement	Ghana Score	Berlin Score
1	The chatbot's personality was realistic and engaging.	4.0 ± 0.7	3.7 ± 1.1
2	The chatbot seemed too robotic.	2.8 ± 1.0	3.2 ± 1.2
3	The chatbot was welcoming during initial setup.	4.1 ± 0.6	4.2 ± 1.2
4	The chatbot seemed very unfriendly.	1.6 ± 0.8	1.4 ± 0.7
5	The chatbot explained its scope and purpose well.	4.3 ± 0.6	4.1 ± 0.9
6	The chatbot gave no indication of its purpos.e	1.6 ± 1.0	1.6 ± 0.7
7	The chatbot was easy to navigate.	4.5 ± 0.5	4.8 ± 0.4
8	It would be easy to get confused when using the chatbot.	1.8 ± 1.0	1.9 ± 0.8
9	The chatbot understood me well.	3.5 ± 1.0	3.7 ± 1.1
10	The chatbot failed to recognise a lot of my inputs.	2.5 ± 1.2	2.9 ± 1.6
11	The chatbot response were useful, appropriate and informative.	3.9 ± 0.8	3.9 ± 0.9
12	The chatbot responses were irrelevant.	1.6 ± 0.6	1.9 ± 1.1
13	The chatbt coped well with any errors or mistakes.	3.6 ± 1.1	3.6 ± 0.7
14	The chatbot seemed unable to handle any errors.	2.2 ± 0.9	2.3 ± 1.0
15	The chatbot was very easy to use.	4.7 ± 0.5	4.7 ± 0.5
16	The chatbot was very complex.	1.8 ± 0.8	1.2 ± 0.4

Table 4. Level of Agreement to the CUQ Statements at the Ghana and Berlin Workshops (Results Expressed as Mean +/- SD)

human-like design, the opinions of the participants differed. The employees of NBSG were in favour of the emojis, making it feel like chatting on WhatsApp with another human being, as well as the response delays that made the chatbot seem less automated and robotic. On the contrary, the employees of SANBS/WCBS recommended to reduce the emojis and provide prompt answers, because particularly for addressing the youth it is important to speak their language and meet their expectations. Regarding the output of the chatbot, the participants emphasised that the chatbot should ensure that the user is satisfied with the answer to his/her request before asking another question on another topic. Furthermore, they recommended to add additional multimedia content like links to videos, for example. To address this, we already incorporated two more links, one forwarding to a virtual 3D tour, explaining how the blood donation process works with all the steps being involved. The second links to a PDF document containing a cookbook with recipes to prevent iron deficiency, both provided by the website of the WCBS. We additionally included dialogue continuity suggestions in the form of buttons for triggering four related follow-up questions to eleven of our 82 modelled user intents on Google Dialogflow. Overall, the feedback of both focus group workshops was very promising and we received important insights for further research as well as improvements of our artefact.

Conclusion and Outlook

In this work, we have addressed the research question on how to design chatbots for potential blood donors to promote sustainable prosocial behaviour with the aim of increasing blood donations and the number of regular blood donors. We focused on the two African countries Ghana and South Africa, for which the number of blood donations decreased significantly due to COVID-19, starting from an already very low percentage of regular donors within the population.

In order to design our artefact, we applied the design science research framework by Kuechler and Vaishnavi (2008) and grounded our research on the literature of behaviour change models. Building on our previous work, we executed in total two design cycles with the input of experts from three blood services from both countries. We developed three design principles and instantiated them in a chatbot prototype developed with Google Dialogflow providing NLP capabilities for user intent detection and a custom-built web interface for convenient access. The developed chatbot was evaluated in two workshops with participants from South Africa and Ghana that included focus group discussions and online surveys. Overall, the chatbot was perceived as useful and was well-received by the experts, even more so from the Ghanaian blood service. The experts especially saw a benefit of using the chatbot for donor education, confirming DP2. DP2 addresses the biggest challenge regarding blood donor behaviour change derived from theory, i.e., TPB, TTM and the presented persona-based user archetypes, which is especially the case for non- and first-time donors in order to overcome a lack of knowledge and receive education.

While the evaluation was very positive, there are also limitations to our research. So far, the chatbot only fulfils the reactive design requirements and focuses on providing information when asked about it. The proactive DPs (DP1 and DP3) comprising push notifications and reminders that will additionally strengthen the ability of the chatbot to facilitate behaviour change will be addressed in future research and the interplay between the chatbot's design and blood donor behaviour change will be explored over a longer period of time and in the field. Nevertheless, the chatbot in its current form exceeds standard FAQ websites due to the integration of images and links and because it allows more natural and engaging interactions, potentially leading to a closer relationship with the user. Consequently, the current chatbot will predominantly appeal to non-donors and potentially first-time donors. While this can be seen as a limitation, these are also the two most important groups for the two countries with regard to recruiting regular donors.

Even though it turned out to be challenging, we aim to perform an online survey with South African and Ghanaian inhabitants to evaluate and refine the design principles now that the two blood services have confirmed the usabilty of the chatbot. As a next step, we aim to extend our research to a holistic blood donation app that includes the chatbot as a component. This will allow us to include proactive aspects of the chatbot as for example push-notifications about current demand or reminders on upcoming blood donation appointments. In previous research, we started to investigate blood donation apps and chatbots for the use in Germany. Together with the research focused on African countries that we presented in this paper, we want to analyse the differences in design between the countries and from what they are originating (e.g., cultural reasons). This will allow us to derive even more generic design principles and design theories for blood donation IS. As the experts from Ghana see great potential in the chatbot, also for internal use cases, another future research project will deal with the design of a collaborative chatbot for the blood service employees. We will additionally investigate the applicability of a chatbot as an interface to a blood supply chain information and decision support system that is currently developed as part of a funded research project.

Acknowledgement

This research was conducted within the project "Blood Information System for Crisis Intervention and Management" (BISKIT) funded by the German Federal Ministry of Research and Education.

References

Ajzen, I. 1991. "The theory of planned behavior," *Organ Behav Hum Decis Process* (50:2), pp. 179–211.

Amershi, S. et al. 2019. "Guidelines for Human-AI Interaction," in: *Proceedings of the 2019 Conference on Human Factors in Computing Systems,* Glasgow.

Amoyal, N., Robbins, M., Paiva, A., Burditt, C., Kessler, D., and Shaz, B. 2013. "Measuring the Processes of Change for Increasing Blood Donation in Black Adults," *Transfusion* (53:6), pp. 1280–1290.

Asamoah-Akuoko, L., Ullum, H., Appiah, B., Hassall, O. W., Ndanu, T., Adongo, P., and Bates, I. 2021. "Determinants of intention to return to donate blood among first-time blood donors in Ghana," *Vox Sang* (116:3), pp. 324–335.

Bagozzi, R. 1981. "Attitudes, intentions, and behavior: a test of some key hypotheses," *J Pers Soc Psychol* (41), pp. 607–627.

Bagozzi, R. 1982. "A field investigation of causal relations among cognitions, affect, intentions, and behavior," *J Marketing Res* (19), pp. 562–584.

Batis, A. A. and Albarrak, A. 2021. "Preferences and features of a blood donation smartphone app: A multicenter mixed-methods study in Riyadh, Saudi Arabia," *Comput Methods Program Biomed* (1).

Bednall, T. C., Bove, L. L., Cheetham, A., and Murray, A. L. 2013. "A systematic review and meta-analysis of antecedents of blood donation behavior and intentions," *Soc Sci Med* (96), pp. 86–94.

Bhattacherjee, A. and Lin, C. 2015. "A unified model of IT continuance: three complementary perspectives and crossover effects," *Eur J Inf Syst* (24:4), pp. 364–373.

Burditt, C., Robbins, M., Paiva, A., Velicer, W., Koblin, B., and Kessler, D. 2009. "Motivation for blood donation among African Americans: developing measures for stage of change, decisional balance, and self-efficacy constructs," *J Behav Med* (32:5), pp. 429–442.

Canadian Blood Services 2017. *Canadian Blood Services launches new tool to engage young donors*. https://www.blood.ca/en/about-us/media/newsroom/canadian-blood-services-launches-new-tool-engage-young-donors (visited on 24 March 2023).

Charng, H. W., Piliavin, J. A., and Callero, P. 1988. "Role identity and reasoned action in the prediction of repeated behavior," en. *Soc Psychol Q* (51), pp. 303–317.

Dale, R. 2016. "The return of the chatbots," *Natural Language Engineering* (22:5), pp. 811–817.

Diederich, S., Lichtenberg, S., Brendel, A. B., and Trang, S. 2019. "Promoting sustainable mobility beliefs with persuasive and anthropomorphic design: Insights from an experiment with a conversational agent," in: *Proceedings of the 40th International Conference on Information Systems,* Munich.

Feine, J., Gnewuch, U., Morana, S., and Maedche, A. 2019. "A Taxonomy of Social Cues for Conversational Agents," *Int J Hum Comput Stud* (132), pp. 138–161.

Ferguson, E. 1996. "Predictors of future behaviour: A review of the psychological literature on blood donation," *Br J Health Psychol* (1), pp. 287–308.

Ferguson, E. and Bibby, P. A. 2002. "Predicting future blood donor returns: past behavior, intentions, and observer effects," *Health Psychol* (21:5), pp. 513–518.

Ferguson, E. and Chandler, S. 2005. "A stage model of blood donor behaviour: assessing volunteer behaviour," *J Health Psychol* (10:3), pp. 359–372.

Floyd, I. R., Cameron Jones, M., and Twidale, M. B. 2008. "Resolving incommensurable debates: A preliminary identification of persona kinds, attributes, and characteristics," *Artifact* (2), pp. 12–26.

Følstad, A. and Brandtzæg, P. B. 2017. "Chatbots and the new world of HCI," *Interactions* (24), pp. 38–42.

GeoPoll 2020. *Resource: Smartphone Usage and Internet Costs in South Africa*. https://www.geopoll.com/resources/south-africa-smartphone-internet-usage/ (visited on 24 March 2023).

Giles, M. and Cairns, E. 1995. "Blood donation and Ajzen's theory of planned behaviour: an examination of perceived behavioural control," *Br J Soc Psychol* (34), pp. 173–188.

Gnewuch, U., Morana, S., and Maedche, A. 2017. "Towards Designing Cooperative and Social Conversational Agents for Customer Service," in: *Proceedings of the 38th International Conference on Information Systems,* Seoul.

Godin, G., Sheeran, P., Conner, M., Germain, M., Blondeau, D., Gagné, C., Beaulieu, D., and Naccache, H. 2005. "Factors explaining the intention to give blood among the general population," *Vox Sang* (89:3), pp. 140–149.

Godin, G., Conner, M., Sheeran, P., Bélanger-Gravel, A., and Germain, M. 2007. "Determinants of repeated blood donation among new and experienced blood donors," *Transfusion* (47:9), pp. 1607–1615.

Hassenzahl, M. 2001. "The effect of perceived hedonic quality on product appealingness," *J Hum-Comput Int* (13), pp. 481–499.

Hevner, A. 2007. "A three cycle view of design science research," *Scand J Inf Syst* (19:2), pp. 87–92.

Holmes, S., Moorhead, A., Bond, R., Zheng, H., Coates, V., and Mctear, M. 2019. "Usability testing of a healthcare chatbot: Can we use conventional methods to assess conversational user interfaces?," in: *Proceedings of the 31st European Conference on Cognitive Ergonomics,* Belfast, pp. 207–214.

Kuechler, W. and Vaishnavi, V. 2008. "On theory development in design science research: anatomy of a research project," *Eur J Inf Syst* (17:5), pp. 489–504.

Kyle, T. and Silver, L. 2019. *Smartphone Ownership Is Growing Rapidly Around the World, but Not Always Equally*. https://www.pewresearch.org/global/wp-content/uploads/sites/2/2019/02/Pew-Research-Center_Global-Technology-Use-2018_2019-02-05.pdf (visited on 24 March 2023).

Laugwitz, B., Held, T., and Schrepp, M. 2008. "Construction and Evaluation of a User Experience Questionnaire," in: *HCI and Usability for Education and Work. USAB 2008. LNCS,* A. Holzinger (ed.). Berlin, Heidelberg: Springer, pp. 63–76.

Maedche, A., Legner, C., Benlian, A., Berger, B., Gimpel, H., Hess, T., Hinz, O., Morana, S., and Söllner, M. 2019. "AI-Based Digital Assistants. Opportunities, Threats, and Research Perspectives," *Bus Inf Syst Eng* (61:4), pp. 535–544.

Martín, J., Muñoz-Romero, C., and Ábalos, N. 2017. *Improve your chatbot's design*. https://chatbottest.com/ (visited on 24 March 2023).

Masser, B., White, K., Hyde, M., and Terry, D. 2008. "The psychology of blood donation: current research and future directions," *Transfus Med Rev* (22:3), pp. 215–233.

Mayroth, N. 2017. *Meet the chatbot that helps you donate blood*. https://www.thehindu.com/news/cities/mumbai/meet-the-chatbot-that-helps-you-donate-blood/article20735945.ece (visited on 24 March 2023).

Morana, S., Friemel, C., Gnewuch, U., Maedche, A., and Pfeiffer, J. 2017. "Interaktion mit smarten Systemen - Aktueller Stand und zukünftige Entwicklungen im Bereich der Nutzerassistenz," *WuM* (9:5), pp. 42–51.

Müller, H. M. and Reuter-Oppermann, M. 2022a. "A Design Science Approach to Blood Donation Apps," in: *The Transdisciplinary Reach of Design Science Research. DESRIST 2022. LNCS,* A. Drechsler, A. Gerber, and A. Hevner (eds.). Cham: Springer, pp. 221–232.

Müller, H. M. and Reuter-Oppermann, M. 2022b. "Chatblood - Towards designing chatbots for blood donors," in: *Proceedings of the 30th European Conference on Information Systems,* Timisoara.

Müller, H. M. and Reuter-Oppermann, M. 2023. "Persuasive Blood Donation App Design for Individualist and Collectivist Cultures," in: *Design Science Research for a New Society: Society 5.0. DESRIST 2023. LNCS,* A. Gerber and R. Baskerville (eds.). Cham: Springer, pp. 157–172.

Muthivhi, T., Olmsted, M., Park, H., Sha, M., Raju, V., Mokoena, T., Bloch, E., Murphy, E., and Reddy, R. 2015. "Motivators and deterrents to blood donation among Black South Africans: a qualitative analysis of focus group data," *Transfus Med* (25:4), pp. 249–258.

Nass, C. and Moon, Y. 2000. "Machines and Mindlessness: Social Responses to Computers," *J Soc Issues* (56:1), pp. 81–103.

Nass, C., Steuer, J., and Tauber, E. R. 1994. "Computers are Social Actors," in: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems,* Boston, pp. 72–78.

Peffers, K., Rothenberger, M. A., Tuunanen, T., and Vaezi, R. 2012. "Design Science Research Evaluation," in: *Design Science Research in Information Systems. Advances in Theory and Practice. DESRIST 2012. LNCS,* K. Peffers, M. A. Rothenberger, and W. L. Kuechler (eds.). Berlin: Springer, pp. 398–410.

Piliavin, J. A. and Callero, P. L. 1991. *Giving Blood: The Development of an Altruistic Identity,* Baltimore: Johns Hopkins University Press.

Prochaska, J. and DiClemente, C. 1982. "Transtheoretical therapy: toward a more integrative model of change," *Psychother Theory Res Pract* (19:3), pp. 276–288.

Roman, M. K., Bellei, E. A., Biduski, D., Pasqualotti, A., Araujo, C. D. S. R., and Marchi, A. C. B. 2020. ""Hey assistant, how can I become a donor?" The case of a conversational agent designed to engage people in blood donation," *J Biomed Inform* (107).

SANBS 2023. *Blood stock levels*. https://sanbs.org.za/ (visited on 24 March 2023).

The American National Red Cross 2023. *Meet Clara, the Blood Donation Chatbot. Use the power of AI to help you manage your appointment and learn about donations*. https://www.redcrossblood.org/donate-blood/dlp/meet-clara--the-blood-donation-chatbot-.html (visited on 24 March 2023).

Tremblay, M. C., Hevner, A. R., and Berndt, D. J. 2010. "Focus Groups for Artifact Refinement and Evaluation in Design Research," *Commun Assoc Inf Syst* (26:27), pp. 599–618.

Venable, J. R., Pries-Heje, J., and Baskerville, R. 2016. "FEDS: a Framework for Evaluation in Design Science Research," *Eur J Inf Syst* (25), pp. 77–89.

Verhagen, T., Van Nes, J., Feldberg, F., and Van Dolen, W. 2014. "Virtual Customer Service Agents: Using Social Presence and Personalization to Shape Online Service Encounters," *J Comput-Mediat Comm* (19:3), pp. 529–545.

Vermeulen, M. et al. 2019. "Assessment of HIV transfusion transmission risk in South Africa: a 10-year analysis following implementation of individual donation nucleic acid amplification technology testing and donor demographics eligibility changes," *Transfusion* (59), pp. 267–276.

Vries, H., Mudde, A., Dijkstra, A., and Willemsen, M. 1998. "Differential beliefs, perceived social influences, and self-efficacy expectations among smokers in various motivational phases," *Prev Med* (27:5 Pt 1), pp. 681–689.

Weber, P. and Ludwig, T. 2020. "(Non-)Interacting with Conversational Agents: Perceptions and Motivations of Using Chatbots and Voice Assistants," in: *Proceedings of the Conference on Mensch und Computer,* Magdeburg, pp. 321–331.

Weizenbaum, J. 1966. "ELIZA—A Computer Program For the Study of Natural Language Communication Between Man And Machine," *Commun ACM* (9), pp. 36–45.

Zarour, M. and Alharbi, M. 2017. "User experience framework that combines aspects, dimensions, and measurement methods," *Cogent Eng* (4), pp. 1–25.