Putting the Value in VR

How to Systematically and Iteratively Develop a Value-Based VR Application with a Complex Target Group

Hanneke Kip
Centre for eHealth and Wellbeing Research
University of Twente
Enschede, the Netherlands
Department of Research
Stichting Transfore
Deventer, the Netherlands
h.kip@utwente.nl

Saskia M. Kelders
Centre for eHealth and Wellbeing Research
University of Twente
Enschede, the Netherlands
Optentia Research Focus Area
North-West University
Vanderbijlpark, South Africa
s.m.kelders@utwente.nl

Lisette J.E.W.C. van Gemert-Pijnen
Centre for eHealth and Wellbeing Research
University of Twente
Enschede, the Netherlands
j.vangemert-pijnen@utwente.nl

ABSTRACT

In development, implementation and evaluation of eHealth it is essential to account for stakeholders' perspectives, opinions and values, which are statements that specify what stakeholders want to achieve or improve via a technology. The use of values enables developers to systematically include stakeholders' perspectives and the context of use in an eHealth development process. However, there are relatively few papers that explain how to use values in technology development. Consequently, in this paper we show how we formulated values during the multi-method, interdisciplinary and iterative development process of a VR application for a complex setting: forensic mental healthcare. We report the main foundations for these values: the outcomes of an online questionnaire with patients, therapists and other stakeholders (n=146) and interviews with patients and therapists (n=18). We show how a multidisciplinary project team used these qualitative results to formulate and adapt values and create lo-fi prototypes of a VR application. We discuss the importance of a systematic development process with multiple formative evaluations for eHealth and reflect on the role of values within this.

CHI 2019, May 4–9, 2019, Glasgow, Scotland Uk

ACM ISBN 978-1-4503-5970-2/19/05…$15.00
https://doi.org/10.1145/3290605.3300365

CCS CONCEPTS

• **Human-centered computing** → *HCI design and evaluation methods*; • **Applied computing** → *Law, social and behavioral sciences*;

KEYWORDS

eHealth; value-based technology; formative evaluation; forensic mental healthcare; iterative development; virtual reality

ACM Reference Format:
Hanneke Kip, Saskia M. Kelders, and Lisette J.E.W.C. van Gemert-Pijnen. 2019. Putting the Value in VR: How to Systematically and Iteratively Develop a Value-Based VR Application with a Complex Target Group. In *CHI Conference on Human Factors in Computing Systems Proceedings (CHI 2019), May 4–9, 2019, Glasgow, Scotland Uk.* ACM, New York, NY, USA, 12 pages. https://doi.org/10.1145/3290605.3300365

1 INTRODUCTION

For an eHealth technology to be successful in practice, a good fit with the end-users and their context is key [1–3] This is not something that can be achieved from behind a desk: in line with approaches such as user-centered design, researchers and developers should apply a thorough development process that pays close attention to the needs and wishes of stakeholders such as patients and therapists [3–5]. However, potential end-users such as healthcare professionals and patients often lack time and motivation to participate, so a well-planned and structured approach is important because it is impossible to apply endless iterations [6]. In order to achieve this and overcome the aforementioned major challenges, studies that show how this can be done in different contexts are essential. In order to prevent that the incorporation of stakeholder's needs and wishes is non-systematical and non-replicable, more studies that apply a theory-based

development process and show which steps can be undertaken to involve end-users are required. One way to account for the perspectives of stakeholders and their context is by eliciting values: an ideal or interest a stakeholder aspires to or has [3, 6]. Values specify what stakeholders want to achieve or improve via a technology. This study aims to show one approach towards eliciting these values from intended end-users and other stakeholders such as healthcare managers and technology developers. The approach is applied to the development process of a VR application for forensic mental healthcare. In this approach, methods and principles from user-centered design are applied by constantly evaluating with potential end-users and creating and evaluating prototypes and scenarios [4].

Especially in settings with end-users that are hard to reach, motivate and keep involved it is important to conduct a good development process that emphasizes the needs and wishes of these people. It appears to be especially challenging to design for vulnerable populations that may be disadvantaged because of for example mental illness, low literacy, little education, low income or drug dependencies [7, 8]. Especially with vulnerable people that might have difficulties in formulating their opinions, it is important to determine what their values are. Yet, there seem to be very few studies that show how to do this by means of putting their values central. Consequently, this study will focus on a part of the development process of an eHealth technology for such a vulnerable population: forensic psychiatric in- and outpatients.

Treatment of forensic patients takes place at the intersect between law and mental health. One of its main goals is to prevent recidivism of aggressive or sexual offending [9]. Studies show that many forensic patients are unmotivated for their often mandated treatment [10]. Also, forensic patients often suffer from multiple psychiatric disorders [11], are disproportionately poor, unemployed and lower educated [12]. Partly because of the complex and unique nature of this population, existing eHealth technologies cannot be copy-pasted into this domain and are often not as successful as expected [13]. Consequently, more attention should be paid to the development of useful technologies for this target group and accompanying development methods that suit its unique characteristics [14].

A technology that seems to fit the forensic psychiatric patient population is virtual reality (VR). Multiple studies indeed point out the added value of VR for forensic patients [14–16]. A recent qualitative study we conducted showed a broad range of potential advantages of VR for forensic mental healthcare [17]. Important advantages are the emphasis on practicing skills instead of abstract talking, the use of VR as a bridge between a closed therapy room and the outside world and the possibility of exposing patients to offense-related situations to create new insights for therapists and patients themselves. However, these results are broad, abstract directions and are not yet suitable for creating a VR application. The current study builds upon the results of this qualitative study by showing how concrete values for a VR application for Dutch forensic mental healthcare have been elicited by using multiple methods, derived from user-centered design principles. We will show how to formulate values in a systematic yet iterative way for a complex end-user group and context.

2 BACKGROUND AND RELATED WORK

Technology and forensic mental health

A recent systematic review pointed out that technology can be of added value for patients, therapists and the quality of care in forensic mental health in multiple ways [14]. This review included 50 studies that focused on a broad range of technologies, for example web-based interventions, serious games, videoconferencing, social platforms and virtual reality. Based on the advantages that were mentioned in these articles, it was concluded that an important way in which technology can add value to treatment is its ability to allow users to interact with realistic situations taking place in the outside world, as opposed to regular treatment in a therapy room. Furthermore, multiple studies found that therapists and even patients often have a positive attitude towards working with technologies such as VR in treatment, which might increase the usually low treatment motivation of patients [10]. Also, an important benefit of technology is that it can be tailored to individual patient characteristics to overcome a one-size-fits-all approach that does not fit the diverse forensic patient population [14]. A technology which has characteristics that are aligned with these benefits is virtual reality.

Virtual reality

An important characteristic of VR is that patients can enter computer-generated worlds in which real-world sensory visual and auditory perceptions are replaced with virtual ones [18]. This can create a sense of presence, which is the feeling of actually being in a virtual place while being somewhere else [19]. This sense of presence allows for a broad range of possibilities of VR for treatment of psychiatric disorders. Indeed, VR interventions prove to be efficacious and promising forms of psychological treatment in general [20]. While very little is known about VR in forensic mental health, there are many possible advantages [14–17]. For example, inpatients often reside in secured settings for longer periods of time, which makes it difficult for them to practice skills in a realistic way [17]. VR can serve as a bridge between a closed setting and the outside world. Furthermore, VR can

be used to validly assess offenders' behavioral and physiological responses to stimuli, for example by showing virtual young children to pedophiles [21]. Furthermore, VR can be used to overcome practical issues with exposure therapy. To illustrate, inpatients are often not allowed to go outside, which results in difficulties with exposure therapy for fear of heights. Also, it might be helpful to expose addicted patients to drugs, but showing real narcotics for this provides ethical and practical issues [17]. Most of these advantages are potential, since not much research on the use of VR in forensic mental health has been published [14].

Especially when not much is known about a technology within a specific context, a good development process is important to ensure that a VR application is of added value for this context and that it fits the needs and characteristics of the users [3, 22]. However, there is still relatively little knowledge on how to develop VR interventions for healthcare [18], especially in complex settings with vulnerable patients. With regard to technology development in the forensic setting, the earlier mentioned systematic review showed that only a few of the included studies discussed technology development in forensic mental health. The ones who did pointed out the importance of incorporating the opinions, preferences and characteristics of users and recommended iterative development with multiple evaluation cycles [14]. However, the suggestions on development were mostly in general terms and abstract: no examples or specific recommendations on how to do this were provided.

eHealth development

When looking at literature on eHealth development in general, the importance of incorporating stakeholder perspectives and continuous formative evaluations is indeed deemed essential [3, 23–26]. However, most papers recommend this in general terms and do not provide concrete examples on how to operationalize these principles. A study that aims to overcome this gap in knowledge by showing part of a participatory, iterative development process in practice is that of Van Velsen et al. [6]. In this paper, the authors formulate values, attributes and accompanying requirements by means of user expressions. The use of requirements is deemed essential for good eHealth development, mostly to ensure that a to-be-developed technology fits the needs, characteristics and goals of its users and their context [3, 4, 27]. Besides requirements, values that state what stakeholders' ideals and interests are should also be formulated [6]. This should be done to prevent a mismatch between the goal of a technology and problems or needs of stakeholders and their context. Values can serve as a bridge between the stakeholder perspective and the technology. Concrete requirements should be matched with abstract values to ensure that they are aligned with the more abstract goals, problems and interests of stakeholders. However, there is a gap between abstract values and very specific requirements. Consequently, to further specify these conceptual values, more concrete attributes are necessary, which are summaries of the needs or wishes that are spoken out by stakeholders [6]. Van Velsen and colleagues aim to show how to formulate values, attributes and requirements via an approach based on the multidisciplinary CeHRes Roadmap for eHealth development. This Roadmap is grounded by, among other things, principles from user-centered design, requirements engineering and value proposition design [1, 3]. The current study is embedded in the CeHRes Roadmap as well. It is mainly based on its second phase, the value specification. Here, results from the previous contextual inquiry phase and newly collected data are used to formulate values, attributes and requirements to establish what a technology should do and look like in order to be of as much added value for stakeholders as possible [3].

While our approach has similarities with that of Van Velsen et al. [6], we operationalized the principles of the CeHRes Roadmap in different ways. Among other things, we use abstract codes from multiple methods instead of user expressions to base values on. Also, the roles of formative evaluations and involvement of users such as patients and therapists are emphasized more, for example by means of an interdisciplinary project team [3, 28]. Finally, the current study focuses on abstract values instead of concrete requirements to point out the importance of first determining the goals and ideals of stakeholders before starting with design. Only after the abstract values have been mapped, researchers should start to formulate specific requirements. Consequently, this paper aims to show the systematic, iterative and multi-method process that was used to formulate values for a VR application for treatment in forensic mental healthcare.

3 METHODS

Study design and ethics

This paper is based on a larger project, called VooRuit met VR, that took place in a forensic hospital in the east of the Netherlands. The main goal of this project was to develop a VR intervention for forensic mental health, using the CeHRes Roadmap [1, 3]. This project was coordinated by a multidisciplinary project team consisting of patients, therapists and researchers. Multiple research methods in which stakeholders were constantly involved were used, and an iterative process with formative evaluation cycles was applied. In the current paper, an online questionnaire, the value formulation and interviews are presented, see Figure 1 for an overview of the order and interrelationship of these methods. The

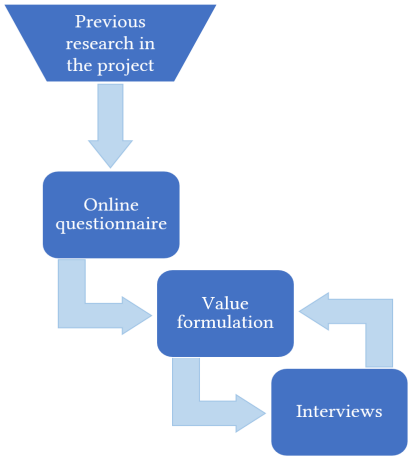

Figure 1: A visualization of the methods of this study and their interrelationships

content of the questionnaire is based on the outcomes of previous research activities of the larger project: focus groups with patients and therapists on possibilities of VR, interviews with patients and therapists on points of improvements in the current situation; and a systematic literature review on VR in treatment of forensic patients.

Since the project took place in practice and involved vulnerable patients, special attention was paid to ethics. Especially the patients and therapists that were part of the multidisciplinary project team played an important role in this. They were actively involved in setting up the studies and paid attention to whether the form and content of the used methods was appropriate for patients and therapists. Among other things, they ensured that the language was understandable and that no sensitive questions were asked. Furthermore, all used materials were elaborately pilot tested with therapists and patients before the actual data collection took place. We also used informed consent forms in all research activities and made sure to inform participants clearly and elaborately on the nature and goal of each study in a way that fits with the patient's abilities. Finally, when including patients in research, we always first contacted involved therapists who knew these patients to ensure that participation would not result in dangerous or uncomfortable situations for patient or researcher.

Online questionnaire

Participants. An online questionnaire was created to gain insight into the preferences and opinions of stakeholders regarding VR in forensic mental health. We aimed to included current and former psychiatric patients, therapists working in forensic mental healthcare and other stakeholders who have a relationship with forensic mental healthcare, such as managers, government employees and technology developers that have experience with designing for forensic mental health. Multiple sampling methods were used to form a convenience sample of these three types of participants.

First, we conducted several activities to include participants from the entire country. The link to the questionnaire was posted on several national websites, newsletters and message boards visited by people working in forensic mental health. Also, we distributed flyers and presented the questionnaire on a national conference on forensic mental health, and made use of the national professional networks of members of the project team. Second, additional sampling activities were undertaken in the forensic hospital where the project group resided. Flyers were distributed in all departments of the hospital and were targeted at employees and in- and outpatients. The link to the questionnaire was posted on the website of the forensic hospitals and was sent to all staff members via e-mail. These activities resulted in a convenience sample of 146 people started the questionnaire, of which 19 were patients (2 female; mean age 41.53; SD = 7.37), 89 therapists working in forensic mental health (62 female; mean age 38.79; SD = 12.51) and the remaining 38 participants were other stakeholders (25 female; mean age 43.95; SD = 14.01).

Materials and procedure. The online questionnaire started with a brief explanation and questions on demographics and background information. Since providing concrete examples has proven to be more suitable for eliciting participants' opinion, as opposed to merely asking them for it [5], six short videos on possible applications of VR in forensic mental health were created by a multidisciplinary project group, consisting of two patients, three therapists, two researchers and a manager. The ideas were based on earlier research that took place in the larger project: the team based the ideas on the outcomes of a previously conducted interview study on the possibilities of VR [17], focus groups and literature research on existing VR applications. All videos took between one and two minutes and contained a brief explanation of the goal of the VR application, the embedment in the existing treatment, an example to illustrate the idea, and an explanation of the desired outcomes. The names of the videos were as followed: (1) Triggers and helpers; (2) Observing and interpreting body language; (3) Body language and the effect on others; (4) Roleplaying in context; (5) Moments of choice; and (6) Crime script. Screenshots of these videos are provided in Figure 2. After watching each video, participants were asked to write down the positive and negative elements of the idea. They were also asked to provide suggestions for improvement. Two quantitative measures were included.

Figure 2: Screenshots of the six videos shown in the online questionnaire: (1) Triggers and helpers; (2) Observing and interpreting body language; (3) Body language and the effect on others; (4) Roleplaying in context; (5) Moments of choice; and (6) Crime script.

First, we used the Personal Involvement Inventory (PII), a 10-item validated questionnaire that measures a person's perceived personal relevance of an idea or product [29]. Since not much is known yet about the PII's suitability for this type of developmental eHealth research [30], participants were also asked to grade the ideas on a scale of 1 to 10. Filling in the questionnaire took most participants between 30 and 40 minutes.

Analysis. Because the data was not normally distributed, non-parametric tests were conducted to detect differences between ideas and the three target groups on the scores PII and grades on a significance level of $\alpha < .05$. The qualitative data was coded inductively by three researchers, using the constant comparative method [31].

Value formulation

Values and accompanying attributes were formulated to map what the added value of the technology should be for patients, healthcare professionals and the forensic treatment, according to stakeholders [3, 6, 32]. First, two researchers independently formulated attributes using the codes of the questionnaires, supplemented with minutes of the project team meetings and literature and code schemes from previous research activities. An attribute is 'a summary of the need or wish that is spoken out by stakeholders' [6]. The lists with attributes from both researchers were compared and combined, resulting in one list with attributes that summarized the preferences of the stakeholders. Next, the two researchers composed groups consisting of attributes that addressed the same type of topic. Based on these groups of attributes, accompanying values were phrased. Values are 'an ideal or interest a stakeholder aspires to or has'. In other words: what do stakeholders want to achieve or improve via the technology? Using the attributes, definitions for the values were written to describe as clearly as possible what the value entailed. The list with values and their definitions was shown to other researchers and the project group in order to receive feedback. Based on their suggestions, the initial table with values, their definitions and accompanying attributes was fine-tuned.

Interviews

During two two-hour project group meetings, the multidisciplinary project team developed a broad idea for a VR application. First, the values, attributes and the positively evaluated elements of the six ideas that were presented in the online questionnaire were discussed, since they served as the foundation for the idea. Next, an initial, broad idea for a VR application was developed means of brainstorming. This process was structured by means of questions such as 'What should the main goal of the idea be?' and 'What elements should it contain?'. Throughout the development process of this idea, the project team continuously checked whether the idea was still consistent with the values. The eventual idea was an elaborated version of one of the six ideas that was presented in the questionnaire: Triggers and Helpers. Positively evaluated elements from the other five ideas were added to the initial Triggers and Helpers idea, resulting in a more elaborate idea, which was also called 'Triggers and Helpers'.

Participants. The idea for the VR application, Triggers and Helpers, was presented to patients and therapists in an interview. The goals were (1) to check in how far the participants' opinions of the idea match the previously formulated values and (2) to determine if any changes to the values and the idea are required to optimally fit the stakeholders' preferences. Via purposive sampling, eight therapists and ten patients of eight different locations of one Dutch forensic hospital were included. The researchers contacted each location of the hospital and aimed to recruit at least one therapist and one patient per location to ensure that perspectives from all types of care were included. This resulted in a sample of eight therapists, of which five were male, four worked in an inpatient clinic, and four worked mainly with outpatients. Ten male patients were interviewed, of which five were outpatients and five were inpatients.

Materials and procedure. Semi-structured interviews of approximately 20 minutes were conducted with the 18 participants. To ensure that the idea and its application in treatment were as clear as possible to the participants, it was presented via a lo-fi prototype and a scenario [4]. In the scenario, the researcher provided an example of how the VR application could be used in treatment of an imaginary patient to identify triggers and train helpers. Here, triggers refer to stimuli that elicit undesired behavior in patients, while helpers are activities that can support patients in successfully dealing with these triggers. The scenario revolved around a fictional forensic psychiatric inpatient with an alcohol addiction, a bipolar disorder and aggression regulation problems. After four years in forensic care, the patient still had no insight in what triggers his aggressive outbursts. The VR application can be used to look for triggers and effective helpers in a personalized way. Via a dashboard with multiple options, the patient and therapist can choose a personally meaningful environment, for example a bar, living room or street. After that, personalized virtual avatars and their emotions can be chosen, such as an angry, broadly built male or an attractive female. Also, visual and auditive stimuli such as alcohol, barking dogs or loud background noises can be picked from a predetermined list. Once the virtual environment is created, the therapist can play the virtual person via a voice morphing microphone to interact with the patient in a virtual role-play. During this role-play the patient and therapist can look for triggers that elicit undesired responses in the patient. In the virtual environment, the rn43 is able to practice skills to deal with triggers, for example counting to ten or looking for distraction on his or her phone. The environment and stimuli can constantly be adapted to better fit the needs of the patient.

To investigate whether this idea matches the stakeholders' needs and wishes, a semi-structured interview with open-ended questions was conducted with the 18 participants. The interview consisted of two parts. First, to check whether the participants would want to use the VR application, open-ended questions based on an adapted version of the Technology Acceptance Model (TAM [33]) were asked. The TAM was used because of its usefulness in creating a a general overview of the cognitions that influence an individual's intention to use a technology [34], in this case the idea for the VR application. Questions were asked to get insight into the participant's attitude towards the idea, the potential perceived usefulness, the potential perceived ease of use in treatment, and the intention to use the to-be-developed VR application. However, studies show that the TAM lacks a focus on characteristics of the person and the context and thus does not provide enough information for a development team in formulating a broad set of values that also address the people and context [34, 35]. In order to obtain this information, the second part of the interview addressed the participant's ideas on the added value of the VR application for the forensic context, patients and therapists. Also, suggestions on the idea itself and its future use in practice were discussed.

Analysis. The interviews were transcribed verbatim and deductively coded by two researchers. First, the elements of the adapted version of the TAM were used as codes to gain insight into intentions and attitudes towards the technology. Second, the interviews were coded with the values that were formulated in a previous phase to check if the participants' opinions and thoughts on the technology were consistent with these values. If necessary, adaptations to the values and the idea for the VR application were made based on the interviews.

4 RESULTS

Online questionnaire

Wilcoxon signed-rank tests showed no significant differences between the PII scores and grades for the different ideas. The qualitative answers to the three open questions provided more insight into the stakeholder perspective. Table 1 shows the outcomes of the inductive qualitative analysis. Based on the questions, three main categories were created: positive remarks, negative remarks and suggestions on VR in forensic mental health. Each category contained the same main codes, but the sub codes differed. These codes were created to serve as the foundation for the values.

Value formulation

In total, 43 attributes were identified, each based on the analyses of at least one executed research activity within the larger project, for example interviews conducted in an earlier phase of this project, a literature review on VR in treatment of forensic patients, previously held focus groups or the questionnaire. The attributes stated what stakeholders found important and relevant in the to-be-developed VR application with regard to a broad range of topics, for example content, usability and use in practice. Several examples of attributes are provided in Table 2. By clustering these relatively concrete attributes by topic, 13 abstract values were created. These values provide insight into what the technology needs to achieve or improve in treatment, in the forensic hospital, and in individual patients and therapists. An example of the relationship between values, attributes and their foundations can be found in Table 2.

Table 1: The codes of the answers to the three open questions, structured by the five main codes that were identified in all three questions

Judgement	Therapist/treatment	Patient	Content	Practical
Positive remarks				
Positive judgement without explanation	Good addition to therapy; practicing safely; good exercise; insight into behavior patient; cause for conversation; good addition to current treatment	Improvement (future) behavior; insight own behavior; insight behavior others; treatment motivation; suitable for specific target group; supporting in reliving	Realism of behavior; personalization of persons; personalization of scenario; personalization of environment	Visual realism; new, different technology
Negative remarks				
Negative judgement without substantiation	Not or difficult to use in treatment; not of added value to current treatment	Dishonesty about effect; dishonesty about input scenario; not suitable for specific group; no effect; elicitation negative feelings	Not enough options for adaptation of persons, scenarios and environment; not generalizable to real life; important element is lacking; no realistic behavior; use of VR is not necessary to reach same goal	No visual realism; hard to use; high costs; takes time during treatment; takes time to learn to use
Suggestions				
Negative or positive recommendation about introduction in practice	Apply for specific type of treatment; good training of therapist; input for VR scenario should be well-prepared	Check whether VR is suitable for patient; co-creation of scenario with patient; apply to specific target group	Combine ideas; use realistic situations; change an aspect of the application	Combine with other technologies; improve visual realism; offer many choices in persons, scenarios & environments; context use VR; posture use VR; constant development

Interviews

Overall, 17 out of 18 participants had a positive attitude towards the idea. In Table 3, the number of interviews in which a specific value was mentioned is provided.

Based on the interviews, it was decided to split the initial value 'easy to use within treatment' into a content and practice-related value. All other values could be retrieved in the evaluations of the participants and no new values were identified, so overall, the values appeared to be a good reflection of the opinion of the users. By using the values to code the data, several important points of attention arose. First of all, practical issues such as ease of use, suitability for specific patients and affordability were seen as potential problems by participants. Implementation appeared to be important. Second, several values were mentioned by almost all participants: personalization, realistic skills training, gaining insight and new possibilities. This might imply that these values are especially important for the participants and thus might have to receive additional attention throughout the development, implementation and evaluation process.

5 DISCUSSION

Applying a systematic approach

This study showed how we formulated values for a VR application for a complex setting: forensic mental healthcare. Multiple methods based on principles of user-centered design were used: we created lo-fi prototypes and scenarios and asked stakeholders to evaluate them [4]. A pitfall of such a development approach is that it might become 'messy' because many different types of data are collected and results can be conflicting. In order to structure the development and decision-making process, a multidisciplinary project team coordinated the project, which is in line with recommendations on eHealth development [3, 28, 36, 37]. Also, to deal with the large amounts of data, this project's approach consisted of systematic and structured research activities. Yet, the approach was also iterative with continuous formative evaluation cycles that allowed for continuous adaptation of ideas and products, which is essential for good eHealth development [3, 23, 26]. In each step, we used a similar cycle to operationalize this structured yet iterative approach, which is visualized in Figure 3. First, we created or updated an idea and made it tangible via methods such as prototyping or

Table 2: Two examples of a value, accompanying attributes, and their foundations.

Foundation	Attributes
Value: Skills training/practice	Stakeholders think it is important that...
Codes questionnaire	• the patient can acquire or improve daily living skills
Good exercise; Safe practicing; Improvement (future) behavior	• the patient can acquire or improve social skills
Literature	•the patient can acquire or improve emotion regulation skills
Fromberger, Jordan & Muller (2014); Renaud et al. (2014)	• interaction with a virtual other, played by the therapist, is possible
Project team	• behavior of virtual others is realistic
Minutes meeting 8 & 10	
Codes interviews previous phase	
Training daily living skills; Training social skills; Training emotion regulation skills; Skills training in context	
Value: Fit with patient	Stakeholders think it is important that...
Codes questionnaire	• the content/storyline of the VR application can be tailored to the needs and characteristics of an individual patient
Personalization of persons; Scenarios; & Environment; Realism of behavior; Not suitable for specific group; Not enough options for adaptation of persons, scenarios and environment; Offer many choices in persons, scenarios & environments	• the appearance of virtual persons can be adapted to the needs and characteristics of an individual patient
Literature	• the virtual environment can be adapted to the needs and characteristics of an individual patient
Renaud et al., 2014	• the VR application contains a broad range of options for type of storyline, environment and virtual persons
Project team	• the patient can be exposed to personally meaningful, emotion-eliciting stimuli in the VR application
Minutes meeting 8 & 10	• The VR application is not used in target groups for which it is unsuitable
Codes interviews previous phase [1]	
Exposure to emotion-eliciting stimuli	

Table 3: The codes of the answers to the three open questions, structured by the five main codes that were identified in all three questions.

Values	Potential value			Points of attention		
	Total (n=18)	Ther. (n=8)	Pat. (n=10)	Total (n=18)	Ther. (n=8)	Pat. (n=10)
Aim application						
Fit with patient	16	8	8	4	3	1
Bridge between therapy room and outside world	14	7	7	4	2	2
Insight into behavior, thoughts and feelings	14	8	6	2	1	1
Improvement of skills	7	5	2	0	0	0
Generalization of skills to daily life	5	4	1	2	1	1
Safety	1	1	0	5	2	3
Treatment motivation	1	0	1	6	2	4
Embedment in treatment						
Unique addition to current treatment	13	6	7	1	0	1
Cooperation between patient and therapist	11	6	5	2	1	1
Content fits well with current treatment	6	5	1	2	2	0
Practically easy to use in current treatment	5	2	3	8	7	1
Use application in practice						
Constant adaptation of application	5	3	2	0	0	0
Widely applicable	1	1	0	9	4	5
Affordability	0	0	0	2	1	1

scenarios, e.g. the six videos and the scenario of the idea that was used in the interviews. Second, the idea was presented to participants using a method that suits their characteristics. We used a questionnaire with brief videos since we expected that participants would prefer to watch videos instead of reading about ideas, and used relatively short interviews. Third, we thoroughly and systemically analyzed the data by first using an inductive approach to stay close to the data

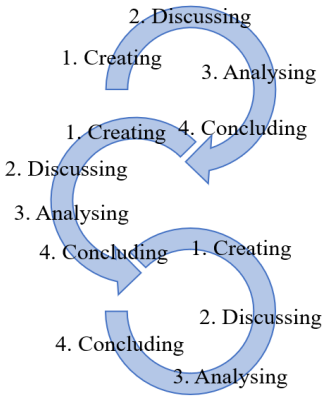

Figure 3: An overview of the systematic and iterative development process of this study

and then a deductive approach to study in how far the values matched the stakeholders' opinions. Fourth, the outcomes of the analyses were used to improve or update the idea for the VR application, the values or other development products such as requirements. In this study, we used the outcomes of the questionnaire to create first versions of attributes and values and combined the six ideas into one. The results of the interviews were used to update the values and their definitions and to make minor improvements to the idea. Because of our positive experiences with this systematic approach, we recommend that in following steps and similar research, this research cycle is used to ensure a structured approach.

Values and formative evaluation

Values state what a technology should achieve according to a stakeholder and can serve as the foundation for a technology throughout the development process [3, 6]. In this study we showed how we formulated values using abstract codes of data from questionnaires and interviews that were conducted in a previous phase of the project [17]. In comparable studies, user expressions from interviews were directly used to ground values and requirements [6, 32, 38]. However, we recommend to also base values and requirements on analyzed results - such as codes - to prevent that researchers get lost in the data. This is especially relevant when much data from different sources is collected, which was the case in the questionnaire, and when using multiple methods. The use of multiple methods was important to create and adapt values and ideas that fit the target group, context and previous results as closely as possible. Central in our process was formative evaluation, which refers to continuous evaluation activities to ensure that there is a fit between the development activities and the stakeholder perspective [3]. Multiple papers on eHealth development acknowledge the importance of such an approach with multiple feedback and evaluation loops [23, 25, 26]. Formative evaluation is an important principle that underlies the four-step cycle that we used multiple times. Ideas that are created should first of all be based on previous results and checked with stakeholders by means of suitable research methods. Second, researchers should make sure that the idea still matches the previously created values. Finally, the results of evaluations of ideas with stakeholders should be used to improve or adapt the idea and, if necessary, the values. A systematic approach with multiple methods supports developers in conducting good formative evaluation activities.

Using values for evaluation

Values can serve as foundations for the design of a technology, but they can also play an important role in implementation and evaluation of eHealth. eHealth implementation is not a separate activity, but is relevant from the beginning of the development process [1, 3, 39]. Consequently, researchers should pay attention to potential implementation issues from the start in order to account for them as soon as possible. One way to do this is by using values. For example, in the interviews, participants made critical remarks or had questions on the value 'practically easy to use'. This implies that the development team should pay close attention to the usability of the VR application before and during implementation since apparently, this value was not apparent enough for stakeholders, which might hinder implementation. Values can also structure evaluation. Since values state what a technology has to achieve, improve or add, it makes sense to base concrete evaluation goals on values [3]. To illustrate: the value 'treatment motivation' states that the VR application should be appealing for patients and increase their motivation to participate in treatment. This suggests that treatment motivation has to be accounted for in the evaluation of the VR application. Another example is the value 'widely applicable': researchers should evaluate whether the VR application is being used in a wide range of patients instead of only a specific group. To our knowledge, there is no research which translates values into evaluation goals and accompanying methods for summative evaluation, so future studies should pay attention to the role of values in evaluation. In order to ensure that development, implementation and evaluation are coherent and intertwined, we recommend that suitable development models are used, for example the CeHRes Roadmap, which is partly based on value proposition design and user-centered design [1, 3].

Dealing with conflicting values

Good eHealth technology closely fits the stakeholder perspective, but this does not mean that all of their needs and

wishes have to be incorporated without question: conflicting values or inconsistencies with existing knowledge might occur [3]. This became apparent in the questionnaire, which showed that while some participants had a very favorable opinion towards one of the six ideas, other participants evaluated the same idea more negatively. Also, the interviews showed that some stakeholders believed that the VR application should be applicable to a broad range of patients, but a majority of the participants questioned if this was possible. There also were conflicts between the stakeholder perspective and scientific knowledge. The results of the questionnaire showed that according to multiple stakeholders, the VR application should look as realistic as possible, comparable to movies. However, animated VR applications are very capable of inducing a sense of presence, which is the feeling of actually being somewhere [40]. On top of that, studies hypothesize that if an animated environment looks too realistic, inconsistencies that disturb believability might arise [41], so an extremely high level of visual realism does not necessarily cause a higher sense of presence. The current study dealt with these types of conflicting values via the multidisciplinary project team that was responsible for decision-making. The team critically discussed conflicting values during meetings and decided how to deal with this. Based on this, we can conclude that developers always have to keep a critical eye on the stakeholders' suggestions to check if they are consistent with scientific knowledge and whether there are any conflicting values. Inconsistencies should be discussed by a multidisciplinary team with a broad range of expertise, ideally also including a member with technical knowledge.

Strenghts and limitations

An important strength of this study was the thorough analysis of the qualitative data [31]. However, these data might have been influenced by a selection bias during convenience sampling. Since only participants with an interest in VR might have participated, the perspective of stakeholders with no interest or an negative attitudes towards VR might be underrepresented. Also, most questionnaire participants and all interview participants were aligned with one forensic hospital in the Netherlands. These two matters might have influenced the generalizability of the results. However, the main goal of both methods was not to paint a generalizable picture, but to provide input for the development of VR. Also, the results were constantly checked by the project team as critically as possible, so we expect that this will not influence the quality of the design. Another issue related to sampling centers around involving patients. Especially in the questionnaire study, it became clear that including patients was challenging, despite our efforts to tailor the content and sampling materials to their characteristics. One explanation is that the questionnaire was too long: watching the videos and answering the open and closed questions often took participating patients over 40 minutes. Especially when working with vulnerable forensic populations who are not always intrinsically motivated to participate, might have difficulties with reading, or have short attention spans [10–12], it seems to be important to keep the data collection as brief as possible [8]. It was indeed easier to recruit patients to participate for the interviews, which only took about 15 to 20 minutes. Furthermore, while we allotted five VR glasses under the participants to reward participating, it might be that this type of rewarding was not tangible enough for the patients. Dugas et al. [8] indeed suggest that participating should have a direct, personally relevant reward for patients. A final limitation is that, while other stakeholders such as developers and healthcare managers participated in the questionnaire, they did not provide their opinions on the idea in the interviews. The values of these types of stakeholders might have differed from those of patients and therapists, so important information could have been missed. However, in following steps, the perspective of other stakeholders will again be incorporated.

Conclusion

Based on this study, we can conclude that values cannot be drawn up from scratch: they have to be based on the results of multiple methods and should be specified via attributes and, later, requirements. This requires a systematic approach to prevent researchers from getting lost in the process and data. First, creating or improving an idea; second, evaluating the idea with stakeholders; third, thoroughly analyzing the data and fourth, making adaptations or improvements to the idea, values or other products. An iterative approach during which values can be constantly fine-tuned is pivotal, mainly to ensure that the values are a good reflection of the stakeholders' opinion. Value elicitation is not a separate step: values should be intertwined throughout and form the foundations of the development, implementation and evaluation of eHealth technology.

ACKNOWLEDGMENTS

Funding for this study was provided by Stichting Vrienden van Oldenkotte. We would like to thank the members of the 'VooRuit met VR' project: Dirk Dijkslag, Yvonne Bouman, Kirby Weerink, Ron Voorhuis, Jakob Visser, Kevin Krimmel and Anne Marike Halma. We thank Ines Brüninghoff and Ankie Kuiper, who helped with the analyses of the questionnaire and interview data.

REFERENCES

[1] Julia EWC van Gemert-Pijnen, Nicol Nijland, Maarten van Limburg, Hans C Ossebaard, Saskia M Kelders, Gunther Eysenbach, and Erwin R Seydel. A holistic framework to improve the uptake and impact of ehealth technologies. *Journal of medical Internet research*, 13(4), 2011.

[2] Russell E Glasgow, Siobhan M Phillips, and Michael A Sanchez. Implementation science approaches for integrating ehealth research into practice and policy. *International journal of medical informatics*, 83(7):e1–e11, 2014.

[3] Hanneke Kip and Lisette JEWC van Gemert-Pijnen. *Holistic development of eHealth technology*, pages 151–186. Routledge, 2018.

[4] Catherine Burns. *Human-centred design*, book section 10. Routledge, Abingdon, 2018.

[5] Nienke Beerlage-de Jong, Jobke Wentzel, Ron Hendrix, and Lisette van Gemert-Pijnen. The value of participatory development to support antimicrobial stewardship with a clinical decision support system. *American journal of infection control*, 45(4):365–371, 2017.

[6] Lex Van Velsen, Jobke Wentzel, and Julia E. W. C. Van Gemert-Pijnen. Designing ehealth that matters via a multidisciplinary requirements development approach. *JMIR Research Protocols*, 2(1):e21, 2013.

[7] Jacquelyn H Flaskerud and Betty J Winslow. Conceptualizing vulnerable populations health-related research. *Nursing research*, 47(2):69–78, 1998.

[8] MichÃĺle Dugas, Marie-ÃĹve Trottier, Selma Chipenda Dansokho, Gratianne Vaisson, Thierry Provencher, Heather Colquhoun, Maman Joyce Dogba, Sophie DupÃĭrÃĭ, Angela Fagerlin, and Anik MC Giguere. Involving members of vulnerable populations in the development of patient decision aids: a mixed methods sequential explanatory study. *BMC medical informatics and decision making*, 17(1):12, 2017.

[9] Paul E Mullen. Forensic mental health. *The British journal of psychiatry*, 176(4):307–311, 2000.

[10] Klaus H Drieschner and Anne Boomsma. The treatment motivation scales for forensic outpatient treatment (tms-f) construction and psychometric evaluation. *Assessment*, 15(2):224–241, 2008.

[11] Kris R Goethals, Ellen CW Vorstenbosch, and Hjalmar JC van Marle. Diagnostic comorbidity in psychotic offenders and their criminal history: A review of the literature. 2008.

[12] Elizabeth Greenberg, Eric Dunleavy, and Mark Kutner. Literacy behind bars: Results from the 2003 national assessment of adult literacy prison survey. nces 2007-473. *National Center for Education Statistics*, 2007.

[13] Joyce Bierbooms, Yvonne Bouman, Dirk Dijkslag, Rob Kimpen, JosÃĭ Muller, and Rosemarie Wieske. Do's en don'ts van e-health in de forensische ggz. *Kwaliteit Forensische Zorg (KFZ)*, 2015.

[14] Hanneke Kip, Yvonne HA Bouman, Saskia M Kelders, and Lisette JEWC van Gemert-Pijnen. ehealth in treatment of offenders in forensic mental health: a review of the current state. *Frontiers in psychiatry*, 9:42, 2018.

[15] P Fromberger, K Jordan, and JL Müller. Anwendung virtueller realitäten in der forensischen psychiatrie. *Der Nervenarzt*, 85(3):298–303, 2014.

[16] Massil Benbouriche, Kevin Nolet, Dominique Trottier, and Patrice Renaud. Virtual reality applications in forensic psychiatry. In *Proceedings of the 2014 Virtual Reality International Conference*, page 7. ACM, 2014.

[17] Hanneke Kip, Saskia M Kelders, Kirby Weerink, Ankie Kuiper, Ines Brüninghoff, Yvonne H A Bouman, Dirk Dijkslag, and Lisette JEWC van Gemert-Pijnen. Identifying the added value of virtual reality for treatment in forensic mental health: a scenario-based, qualitative approach. Submitted.

[18] Daniel Freeman, Sarah Reeve, A Robinson, Anke Ehlers, David Clark, Bernhard Spanlang, and Mel Slater. Virtual reality in the assessment, understanding, and treatment of mental health disorders. *Psychological medicine*, 47(14):2393–2400, 2017.

[19] Julia Diemer, Georg W Alpers, Henrik M Peperkorn, Youssef Shiban, and Andreas Mühlberger. The impact of perception and presence on emotional reactions: a review of research in virtual reality. *Frontiers in psychology*, 6:26, 2015.

[20] Wesley A Turner and Leanne M Casey. Outcomes associated with virtual reality in psychological interventions: where are we now? *Clinical psychology review*, 34(8):634–644, 2014.

[21] Patrice Renaud, Dominique Trottier, Joanne-Lucine Rouleau, Mathieu Goyette, Chantal Saumur, Tarik Boukhalfi, and Stephane Bouchard. Using immersive virtual reality and anatomically correct computer-generated characters in the forensic assessment of deviant sexual preferences. *Virtual Reality*, 18(1):37–47, 2014.

[22] Byeol Kim, Warren Schwartz, Danny Catacora, and Monifa Vaughn-Cooke. Virtual reality behavioral therapy. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 60(1):356–360, 2016.

[23] Susan Michie, Lucy Yardley, Robert West, Kevin Patrick, and Felix Greaves. Developing and evaluating digital interventions to promote behavior change in health and health care: recommendations resulting from an international workshop. *Journal of medical Internet research*, 19(6), 2017.

[24] Lucy Yardley, Leanne Morrison, Katherine Bradbury, and Ingrid Muller. The person-based approach to intervention development: application to digital health-related behavior change interventions. *Journal of medical Internet research*, 17(1), 2015.

[25] Eric B Hekler, Predrag Klasnja, William T Riley, Matthew P Buman, Jennifer Huberty, Daniel E Rivera, and Cesar A Martin. Agile science: creating useful products for behavior change in the real world. *Translational behavioral medicine*, 6(2):317–328, 2016.

[26] Kevin Patrick, Eric B Hekler, Deborah Estrin, David C Mohr, Heleen Riper, David Crane, Job Godino, and William T Riley. The pace of technologic change: implications for digital health behavior intervention research, 2016.

[27] Hossein Saiedian and Robert Dale. Requirements engineering: making the connection between the software developer and customer. *Information and Software Technology*, 42(6):419–428, 2000.

[28] Sue S Feldman, Benjamin L Schooley, and Grishma P Bhavsar. Health information exchange implementation: lessons learned and critical success factors from a case study. *JMIR medical informatics*, 2(2), 2014.

[29] Judith Lynne Zaichkowsky. The personal involvement inventory: Reduction, revision, and application to advertising. *Journal of advertising*, 23(4):59–70, 1994.

[30] Saskia M Kelders. Involvement as a working mechanism for persuasive technology. In *International Conference on Persuasive Technology*, pages 3–14. Springer, 2015.

[31] Hennie Boeije. A purposeful approach to the constant comparative method in the analysis of qualitative interviews. *Quality and quantity*, 36(4):391–409, 2002.

[32] Jobke Wentzel, Lex van Velsen, Maarten van Limburg, Nienke de Jong, Joyce Karreman, Ron Hendrix, et al. Participatory ehealth development to support nurses in antimicrobial stewardship. *BMC medical informatics and decision making*, 14(1):45, 2014.

[33] Viswanath Venkatesh and Fred D Davis. A theoretical extension of the technology acceptance model: Four longitudinal field studies. *Management science*, 46(2):186–204, 2000.

[34] Lex van Velsen, Mirka Evers, Cristian-Dan Bara, Harm Op den Akker, Simone Boerema, and Hermie Hermens. Understanding the acceptance of an ehealth technology in the early stages of development: An end-user walkthrough approach and two case studies. *JMIR Formative Research*, 2(1):e10474, 2018.

[35] Paul Legris, John Ingham, and Pierre Collerette. Why do people use information technology? a critical review of the technology acceptance model. *Information and management*, 40(3):191–204, 2003.

[36] Claudia Pagliari. Design and evaluation in ehealth: challenges and implications for an interdisciplinary field. *Journal of medical Internet research*, 9(2), 2007.

[37] Jeppe Agger Nielsen and Lars Mathiassen. Interpretive flexibility in mobile health: Lessons from a government-sponsored home care program. *Journal of medical Internet research*, 15(10), 2013.

[38] Birgitta Bergvall-Kåreborn and Anna Ståhlbröst. User expressions translated into requirements. *Human Technology: An Interdisciplinary Journal on Humans in ICT Environments*, 2010.

[39] Trisha Greenhalgh, Joseph Wherton, Chrysanthi Papoutsi, Jennifer Lynch, Gemma Hughes, Christine A'Court, Susan Hinder, Nick Fahy, Rob Procter, and Sara Shaw. Beyond adoption: a new framework for theorizing and evaluating nonadoption, abandonment, and challenges to the scale-up, spread, and sustainability of health and care technologies. *Journal of medical Internet research*, 19(11), 2017.

[40] Xueni Pan and Antonia F de C Hamilton. Why and how to use virtual reality to study human social interaction: The challenges of exploring a new research landscape. *British Journal of Psychology*, 2018.

[41] Vinoba Vinayagamoorthy, Andrea Brogni, Marco Gillies, Mel Slater, and Anthony Steed. An investigation of presence response across variations in visual realism. In *The 7th Annual International Presence Workshop*, pages 148–155.