The Grid, Classification of eHealth Applications Towards a Better (re)Design and Evaluation

Saskia M. Akkersdijk*, Saskia M. Kelders†, Julia E. W. C. van Gemert-Pijnen‡ and, Louise M. A. Braakman-Jansen§
Center for eHealth Research & Disease Management
University of Twente,
Enschede, the Netherlands
Email: {s.m.akkersdijk*, s.m.kelders†, j.vangemert-pijnen‡, l.m.a.braakman-jansen§}@utwente.nl

Abstract—eHealth holds a diverse range of successful and unsuccessful applications. It is often unclear why this difference in success appears because eHealth is usually approached as a black box. Additionally, we tend to evaluate non-treatment-like and treatment-like applications the same way. Both approaches focus on searching for effects and outcomes. This results in applications that are wrongfully put away. Placing an application on the continuous dimensions of use-structure and caregiver involvement on the grid presented in this paper helps to make conscious and better decisions when (re)designing and evaluating eHealth applications. Positioning on the grid influences the terms ‘user’ and ‘usage’ but also has implications for the way we can evaluate and (re)design eHealth applications. The grid is a tool to gain insight, facilitate thought processes, and start discussions, and is not meant to be a formal and rigid model. This tool helps making conscious choices in (re)design and evaluation of applications.

Keywords—eHealth; classification; evaluation; design.

I. INTRODUCTION

In eHealth, there is a diverse range of successful and unsuccessful applications and interventions. When observing these applications we observe that applications that represent a form of treatment tend to be more successful than those that are more supportive of nature. Often it is unclear why one application is successful, when the other is not. A reason why we cannot always explain the difference in success is because eHealth is often approached as a black box. With a black box we search for its effects and focus on outcomes. It would be better to examine eHealth technology from a holistic perspective, in which the technology itself also has value, and we focus on the mechanisms behind the success. To find these mechanisms it is necessary to open the black box. An important reason why we observe differences in success is that we evaluate non-treatment-like applications the same way we evaluate the treatment-like interventions, focused on outcome measures or usage numbers. This results in applications that are maybe wrongfully put away because they do not measure up to the measurements they are wrongfully compared to. In this paper we search for a tool to give a better inside in the application, which helps with a (re)design and evaluation of that application.

One of the ways applications are often evaluated is measuring to which extent therapies are followed as intended. This measurement of adherence is one of the primary determinants of success in treatments [1], and overall effectiveness of health systems decreases by poor adherence [1]. Although adherence might be one of the primary determinants of success in therapy there are also examples of applications with a low adherence that are successful. An example of this is QuitNet; a program for smoking cessation [2][3][4]. Adherence to this program is low (23%) [5], but the program can be successful in promoting cessation and preventing relapse [2]. These studies show that it is possible for an eHealth application to have a low adherence but still be successful for a certain group of users.

Studies with eMentalHealth interventions often find a high dose - response relationship (also called a usage - outcome relation). An example of this is the study of Bolier et al. [6]. Donkin et al [7] further explored the usage - outcome relation. The study of Donkin investigates which usage metrics are important in predicting and explaining outcomes for an internet-delivered trail targeting depressive symptoms for those with risk factors for or diagnosis of cardiovascular disease (Cardiovascular Risk E-couch Depression Outcome (CREDO)). Their study shows that there is not always a linear dose - response relationship (also called a usage - outcome relation). An example of this is the study of Bolier et al.

There is a broad range of different eHealth applications and variety in how these applications should be used. These variations can be put on a continuous scale. At one end of the continuum (see Figure 1) we see applications forcing the user to use the intervention in a specified way, for example a fixed use of the modules of an intervention. An example of such an application is the Web-based ‘Living to the full’ intervention. ‘Living to the full’ consists of nine lessons which have to be completed in a specific order in a 12-week period [8]. These applications are often a (web-based) program of a method, course or intervention.

At the other end of the continuum we see applications that leave the usage free, without a strict protocol for each user. QuitNet, the application that was mentioned before for cessation treatment, is such an application. This website offers advice and assistance to quit smoking. Usage frequency of the program and how the program is used is left to the user.
Another important factor that varies among different eHealth applications is the involvement of a caregiver. Some eHealth applications are used in close collaboration between patient and caregiver, others with no involvement of a caregiver at all and all variations in between. Caregiver involvement is often found to be necessary to ensure adherence and increase effects for web-based interventions for people with depression symptoms, or chronic conditions [9][10][11][12][13][14].

Knowing where your application is positioned on these two dimensions can help with (re)designing and evaluating your application. These two dimensions form a grid, and applications can be put somewhere on this grid depending on its usage-structure and caregiver involvement. The positioning of an application on the two dimensions influences the term ‘user’ and ‘usage’ but also has implications for the way we can (re)design and evaluate the application. The aim of this paper is to present a tool to give a better inside in the use-structure, after which we will look at the dimension of caregiver involvement. Finally, we will describe some eHealth applications and their position on the grid.

In Section 2 (The Grid), we will take a closer look at the grid, after which we will discuss implications based on the different positions an application can take on the grid in Section 3 (Implications). We will end this paper with a discussion and conclusion in Section 4 (Discussion and Conclusion).

II. THE GRID

Based on our observations of eHealth applications and the extremes we see, we could classify each application on the following grid based on their characteristics:

```
with caregiver

structured, pre-defined use

without caregiver

free use
```

In this section, we will take a closer look at the two dimensions of the grid. We will first look at the dimension of use-structure, after which we will look at the dimension of caregiver involvement. Finally, we will describe some eHealth applications and their position on the grid.

The dimension of use-structure has at one end of the continuum applications that force the user to use the intervention in a specific way (railroad them). This can be in a specific order, for a specific number of times/lessons, or for a specific duration. These interventions often have a specific end that is known beforehand and are often based on theories about mental health behavior like acceptance and commitment therapy (ACT) or cognitive behavioral therapy (CBT). Because they often find their origin in know theories and therapies they are often more ‘treatment’ like and help deliver a form of short term care. As discussed in the introduction, ‘Living to the full’ is a good example for this end of the continuum. ‘Living to the full’ has been published as a self-help book [15] and is based on mindfulness [16][17] and ACT [18]. The intervention consists of nine lessons which have to be completed in a specific order in a 12-week period. Whether participants worked through a lesson in one session or in multiple sessions was up to them [8].

At the other end applications that leave the usage free without a strict protocol for each user. There is no specific order or duration for which this application should be used, therefore they have no specific end. These free-to-use applications often focus more on support and long term care. As discussed in the introduction, ‘QuitNet’, the application for cessation treatment, is a good example for this end of the continuum. This website offers advice to quit smoking, assistance in setting a quit date, tailored information, assessment of motivation and nicotine dependence, practical counseling (skills training and problem solving), tailored assistants in selecting pharmacotherapies and intra- and extra-treatment social support. How QuitNet is used is completely up to the user.

The vertical dimension represents a form of caregiver involvement, which varies among eHealth applications. For example treatment-driven applications involve caregivers usually, while lifestyle interventions can be used autonomously. Research finds caregiver involvement important, but it is not clear what the dosage and frequency of involvement should be [9][10][11][12][13]. With applications that target people with chronic conditions, there often is some form of caregiver involvement. However, these applications often struggle to find their fit into daily life, and adherence is often low [19]. Users find it difficult to embed these application in their own life, while caregivers struggle to embed them into their daily practice [20]. Nonetheless, caregiver involvement is often found to be necessary to ensure adherence and increase effects for web-based interventions for people with depression symptoms, or chronic conditions [14].

To illustrate the positioning of an eHealth application on the grid, we will now position ‘Living to the full’ on the grid, after which we will describe another application (‘My Health Platform’) and its position on the grid.

As discussed in the introduction, ‘Living to the full’ (LttF) consists of nine lessons which have to be completed in a specific order in a 12-week period. The intervention is used without involvement of a caregiver.
We would place ‘Living to the Full’ at the bottom of the left corner on the grid for the following reasons:

- horizontal dimension: Usage of ‘Living to the Full’ (such as how it is used, how often) is pre-defined. Exactly when (time) a lesson is completed is left to the user. Therefore we would place ‘Living to the Full’ just a small bit to the right out on the horizontal dimension.
- vertical dimension: ‘Living to the full’ is a standalone program without caregiver support, usage is completely left to the user. Therefore we would place ‘Living to the Full’ completely at the bottom of the vertical dimension.

‘My HealthPlatform’ (MHP) is an online platform to support self-care and self-management for people with a chronic illness (e.g. increased cardiovascular risk, COPD, Diabetes mellitus type 2). It is designed to help users keep an overview of and be a director of their own health and lifestyle, alone or in cooperation with a caregiver or expert. In MHP they can monitor their health, find information about their condition, but also use one of the lifestyle coaches. While the usage of most of the platform is unstructured, the coaches follow a 12 week schedule.

We would place ‘My HealthPlatform’ at the right-hand side, in the middle of the vertical dimension on the grid for the following reasons:

- horizontal dimension: Usage of MHP (such as how it is used, how often, and whether or not a coach is used) is left to the user. We would not place MHP completely at the right side, because the coaches do require the user to use them in a specific way and for a predetermined number of weeks.
- vertical dimension: Usage of MHP is mostly left to the user. When MHP is used in cooperation with a caregiver, the caregiver is able to see at home measurements of the user, which provides more insight in the health status of their patient.

You can position MHP on a different position on the grid based on other arguments. In this case, especially the vertical dimension of the grid leaves room for discussion. We would like to emphasize that when we would ask multiple people to position the same application on the grid we are very likely to end up with as many different positions as we asked people. We would like to argue that this is perfectly fine, because the main purpose of the grid is to help you think about certain characteristics of your application and about the implications of the positioning on the grid. In the next section we will talk about some of these implications of the different positions on the grid.

The examples discussed did not have a form of caregiver involvement that in our argumentation would be positioned in the upper half of the grid. An example that would be positioned in the upper left corner of the grid could be an application that only is used in a face-to-face consult with a caregiver in which the ‘patient’ and caregiver together work through a fixed number of modules in a fixed order. An example that would be positioned in the upper right corner of the grid could be an application in which you could consult your caregiver whenever you have question about your health problem.

### III. Implications

Positioning on the grid has several implications for the terms ‘usage’ and ‘user’ and for the (re)design and evaluation of eHealth applications. In this section, we will discuss some of these implications. We will start with the implications on the terms ‘user’ and ‘usage’ after which we will discuss implication for (re)design and evaluation.

#### A. User

Defining when someone is a user is quite clear when you are dealing with applications that are on the left-hand side of the grid. A person that uses the application is a user, and one who does not is not a user. With applications that leave the usage up to the user the way people use the application can vary widely, which leads to a discussion about the term ‘user’
in this context. We will discuss some questions around the term user, after which we will give our view on the answers.

An important question is: when does a person become a user of the application? There are several possible answers to this. We could argue that a person who uses the application is a user, but is there a minimum amount of usage before that person becomes a user, or is 10 seconds enough? And what about someone who does not use the application for a long period of time? Is that person not a user during this period? And could we define certain activities in the application that a person must have done before that person is marked as a user?

For applications that focus on monitoring health or increasing health awareness (mostly positioned on the right-hand side of the grid), we can argue that by only becoming aware of such an application a person could potentially be triggered to become more involved in his own health. This means that, in order to have an effect on a person, it does not automatically require that person to use that application. Is this person then a user? We might argue that this person is not a user of the application because he/she did not interact with it. However, the application could still have an effect. In this case the person is not a user in the most common sense of the word, but due to the effect that the application had it balances on the edge of the definition of ‘user’.

When we consulted the people who used MHP (Figure 3), it became clear that they had their own view on being a user. There were quite a lot of people who had used the application only a couple of times and therefore declined to join several studies (interviews, questionnaires, and usability testing) because they did not see themselves as ‘users’. In their minds their definition of a user involves a certain number of reoccurring visits to the application, entering some monitoring data into the system, or participating in the program of a coach. Because they did not meet their own standards of the term user, they thought they could not participate in the study. This example shows that using the system does not equal being a user, at least not for the people who used it. People might have expectations about the intended usage of an application, it is relevant to communicate the intended use to avoid misunderstanding about the usage.

For evaluation purposes the definition of what we would call a user can focus on several aspects:

1) The percentage of registered users who see themselves as user, could be a measurement for evaluation of an application. The number of registered users who see themselves as user tells you about their involvement with an application and this in turn can show which role the application holds in their lives and whether the application helps them.

2) You do not always know beforehand who will be ‘user’. With evaluations of an application it is important to define which group of people can be defined as user, this group does not always include the groups you thought of before the application was used. For one type of evaluation, questionnaire, or interview, another groups of users might be suitable.

3) An important user that is often forgotten is the caregiver. The caregiver can have his own section in the application where he can see caregiver-specific functionality, he can have its own version of the application, or he can have the same functionality as a patient user. It is important to realize that the caregiver is a user as well, a user with different needs than a ‘patient’. Additionally, both users (the caregiver and the ‘patient’) affect each other and how they use the application, which means that both types of users should be included in the (re)design and all evaluations.

B. Usage

The term ‘usage’ (in the context of an application) can mean a lot of things, such as: How often people return to a website, paths that users follow on a website, how often certain elements on a site are used, etc. This can all be measured by logging user actions with timestamps on a website. How we can use these measurement and what they tell us about the system/application differs between the extremes on the horizontal use-structure dimension of the grid. Measuring the use of a system or application is useful and insightful for both ends of the spectrum, but evaluating this use differs and the implications/interpretations are different.

For ‘railroaded’ application, like ‘Living to the full’ (see Figure 2), usage measurements can tell you much about the applications. ‘Railroaded’ application, positioned at the left-hand side of the grid, often are similar or represent a therapy. The user has to follow the structure within the application, do certain actions in a certain order, and use it for a certain amount for it to be successful. Therefore, we can define ‘normal’, or ‘ideal’ use. We can compare the measured use with the way the application should be used (this can be whether someone completed the application, or the use within an application).

Achieving the goal of the application is not completely dependent on the use (the amount and which parts) for applications which leave the user free (right-hand side of the grid). The duration of usage of these applications is often longer and different situations can be seen than with the use of an ‘railroaded’ application. For applications positioned at the right-hand side of the grid there is no definition of ‘normal’ use, in quantity or in order. This is a contrast with applications positioned at the left-hand side of the grid were the ‘correct’ following of the structure is essential.

Because there is no prescribed use for applications at the right-hand side of the grid, we cannot measure to which extend the measured usage deviates from the optimal use. For example, the measured use of an application positioned at the right-hand side of the grid (like ‘MHP’ in Figure 3) could show users that were dormant for maybe months or years, after which they suddenly used it again. This is
unlikely to happen in an application that is ‘railroaded’. Because we cannot define ‘normal use’, adherence, in which we compare the occurring usage with the optimum usage, cannot be measured for applications on the right-hand side of the grid. As mentioned before, applications positioned on the left-hand side often are more treatment like, in contrast with application on the right-hand side that are more supportive and longer term. This leads to a different type of ‘lessons’ and therefore a different kind of use. The same difference also makes for different characteristics of the user (e.g. in type of motivation), which also leads to different use.

The occurring usage and the use-structure of an application go together. When your application is ‘railroaded’ and positioned at the left-hand side of the grid users have all a similar usage pattern, while with an application that leaves the use up to the user the occurring usage patterns can vary greatly.

Even though measurements like adherence are not really suitable for applications that are positioned at the right-hand side of the grid, usage measurements can still be very valuable. These measurements can tell you much about the interaction with the application, which parts are used most, which parts are often used subsequent of each other, after which part do they stop, etc. Knowing more about the interaction with the application is valuable for improving applications, but can also be valuable for finding mechanisms behind application success. Combining usage measurements with use context (what triggered the session) can be used to find a better fit of the content to the context, or improve interaction with the application. By improving the system, and better tuning it to the needs of the users (based on context en measured usage) we can increase the effect of applications.

Finally, when we are looking at the usage of an application we should not forget to observe the usage of the application by the caregiver. Caregivers play an important role in the usage of an application by their ‘patients’ because their usage can be driven by input of said caregiver. When a caregiver does not work with the application as intended or adequate this will influence the usage of the ‘patient’ user as well. When the application is mend to be used with a form of caregiver involvement and the caregiver is less involved than the ‘patient’ expects, the ‘patient’ user will experience less added value of the application.

C. Implications on (re)Design

With an existing application, the grid can be used during the redesign process. Determining the current position of the application on the grid based on the characteristics of the application can help you to reflect on your current application by facilitating the thought process about your applications and its characteristic. A first step in the redesign process is to reflect whether it is feasible to reach the objectives of the application from its current position; is it possible to accomplish the goal of the application with this position or is the position of the application on the grid not suitable for the goal of the application. The second step is to determine if the current position is the best, or if there are better alternative positions. When the current position and the desired position are known, the next step is to identify their differences. Knowing these differences, it is then possible to determine if the application should be changed and can give a indication about how the application should be changed.

Of course you can also use the grid when designing an application from scratch. The grid can help facilitate the thought process of design choices and their effects to make a better conscious decisions.

D. Implications on Evaluation

Implications on the use-structure dimension of an application are about evaluation procedures. ‘Railroaded’ applications can be evaluated by measuring the usage and comparing it to the optimum usage in contrast to applications on the right-hand side of the grid were usage can vary widely. Because usage can vary so widely it is also harder to link measured effects to a specific element of the application. Applications on the right-hand side of the grid are less feasible to evaluate with an Randomized Controlled Trial (RCT), because they are used for a much longer time, which makes it difficult to keep the circumstances constant. Secondly, for applications on the right-hand side of the grid it can be harder to find changes in measurements like quality of life, because they target larger groups and the changes they accomplish are often small. This does not imply that these changes are unimportant. It would be more suitable to evaluate applications that are positioned on the right-hand side of the grid on processes rather than on effects. While applications that are positioned on the left-hand side of the grid are easier to evaluate on effects, because they have a fixed setting and use-time.

IV. DISCUSSION AND CONCLUSION

There are different ways to classify eHealth that provide an overview of eHealth, such as device driven, based on the medium the technology uses (web-based, mobile apps, etc.), context-of-care driven (eCare, eTherapy, eAppointment, ePrevention, etc.), or actor driven (based on the interaction between the actors of such a system). The grid we propose is not meant to replace these classifications, because they provide an overview that our grid does not provide. However, our grid serves as an extension of these. The different classifications mentioned above serve a different need, while they did not serve our need for a simple way to have some guiding when (re)designing and evaluating eHealth applications. We were looking for a better way to help make a conscious choice in order to find a better fit (in (re)design and evaluation). Positioning of an eHealth application on the grid helps to become more aware of implications this has (as discussed in Section III).

Based on the position of an application on the grid we discussed that the term ‘user’ can include a different group of people. Those who ‘use’ an application do not always perceive themselves as a ‘user’, because they have expectations...
about the intended usage of an intervention. It is relevant to communicate the intended use to avoid misunderstanding about the usage. The percentage of users who see themselves as a user might be an additional measurement for evaluating an application, because it includes values about involvement with the application. An important group of ‘users’ that is often forgotten are the caretakers. They often also use the application, and their use or their communication about the application influences the use of the application by their patients.

Adherence is an important measurement for application positioned on the left-hand side of the application. For these applications, we can define ‘normal’ or intended use. Because ‘normal’ or intended use is often a lot harder to define for applications positioned on the right-hand side of the grid (there often is no prescribed use) and usage patterns can vary widely, the measurement of adherence is not suitable. However, usage measurements can be valuable for process evaluations and improvement of the application.

The grid can help with the (re)design process by gaining more insight facilitated by the thought process needed for placement of the application on the grid. It is important to think through whether the intended (or current) position of the application is suitable for the goal of the application, or that another position might be better.

Positioning of an application on the grid (left versus right, and with or without caretaker involvement) influences which sorts of evaluations are suitable. Evaluations can be focused on process or effects and positioning on the left-hand of the grid are more suitable for effects evaluations than positions at the right-hand side of the grid.

The grid is a tool to gain insight, facilitate thought processes, and start discussions, and is not meant to be a formal and rigid model. This tool helps making conscious choices in (re)design and evaluation of applications.

REFERENCES
