8

Intelligent Mobile Support for Therapy Adherence and Behavior Change

Michel Klein, Nataliya Mogles, Arlette van Wissen

Abstract

Mobile applications have proven to be a promising tool to support people in adhering to their health goals. Although coaching and reminder apps abound, few of them are based on established theories of behavior change. The present work continues where Chapter 2 left off. The main focus is the development of the behavior change support system eMate, which uses the computational model COMBI to create tailored interventions in order to encourage patients to adhere to a healthy lifestyle. The implementation of the system and some preliminary results of its functioning will be discussed.

8.1 Validation of the COMBI Model

The eMate project follows a user-centered design approach, and some formative testing has been done to examine the validity of the COMBI model. This model was presented in Chapter 2 and is depicted in Figure 8.1. Two validation studies were conducted, the first with 40 healthy participants, the second with 17 chronic patients. The main findings and the results of the studies are described in this section.

8.1.1 Study 1

8.1.1.1 Method

A total of 40 healthy subjects participated in the study by filling out an anonymous online questionnaire, which consisted of 42 questions and targeted only adherence to physical exercise. The questionnaire was distributed per e-mail, mostly among staff members of the VU University Amsterdam, but also some other people were included. Participants were assured of the anonymous and voluntary character of the survey. No incentives for participation were offered. The study involved 20 males and 20 females. Age ranged between 21 and 64 (Mean = 33, SD = 11). The questionnaire took approximately 15 minutes to complete.

1The authors are mentioned in alphabetical order and have made a comparable contribution to the article. This chapter will appear as part of Klein, M.C.A., Mogles, N., Wissen, A. van. Intelligent Mobile Support for Therapy Adherence and Behavior Change. Journal of Biomedical Informatics, 2014 (to appear)
All of the measures were in Dutch. The questionnaire measured physical activity behavior and its possible determinants. The physical activity domain was chosen to be the subject of the study over medicine intake and healthy food intake behaviours for two reasons. First, healthy subjects are easier to recruit, and medicine intake is not relevant for everyone. Second, physical exercising measurements available to us had a broad scope and referred to different types of physical exercising, e.g. leisure, home and sport activities.

In total, 14 constructs of the COMBI model were assessed, along with the stage of physical activity behavior (precontemplation, contemplation, preparation, action and maintenance) and physical activity. Three externally validated questionnaires were included in the general questionnaire to obtain values for the constructs in the model: UPCC, PANAS, and SQUASH (see Table 8.2 for a description and the measured constructs). The remaining constructs were measured with the help of one or two questions addressing the determinant of interest. The questions were either formulated as multiple choice recorded on 3, 5 or 6 point Likert-scales, or open. The hypothesis was that there would be positive correlations between all determinants that are connected by causal relations, except of the connection between barriers and self-efficacy that was supposed to be negative.

8.1.1.2 Results

Several sets of bivariate Pearson Product-Moment correlations were performed between the constructs that are closely connected in the model graph. The results of the analysis can be found in Table 8.1. Seven out of the twelve correlations were found to be significant. The strongest significant correlations were found between susceptibility and threat (r = .55, P<0.01), threat and physical activity (r = .50, P<0.01), attitude and physical activity (r = .41, P<0.01), and self-efficacy and physical activity (r = .51, P<0.01). Significant correlations were also found between social norms and attitude (r = .39, P<0.05), mood and physical activity (r = .37, P<0.05) and between cues and physical activity (r = -.38, P<0.05). Two of the significant correlations
Table 8.1: Correlations among the constructs of the COMBI model in physical activity behavior domain based on questionnaire responses (N=40)

<table>
<thead>
<tr>
<th></th>
<th>threat</th>
<th>attitude</th>
<th>self-efficacy</th>
<th>mood</th>
<th>cues</th>
</tr>
</thead>
<tbody>
<tr>
<td>susceptibility</td>
<td>.55**</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>severity</td>
<td>-.03</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pros/cons</td>
<td></td>
<td>.25</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>emotions</td>
<td></td>
<td>.24</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>social norms</td>
<td></td>
<td>.39*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>barriers</td>
<td></td>
<td>-.17</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>skills</td>
<td></td>
<td>.15</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>days physical activity</td>
<td>-.50**</td>
<td>.41**</td>
<td>.51**</td>
<td>.37*</td>
<td>-.38*</td>
</tr>
</tbody>
</table>

* Correlation is significant at 0.05 level (2-tailed).
** Correlation is significant at 0.01 level (2-tailed).

were not according to the expected trends that follow from the causal relations in the model: the negative correlations between threat and days of physical activity and between cues and days of physical activity. The other correlations between the constructs of the model were non-significant, though in line with the expected trends. For instance, according to the model, the direction of correlations between pros/cons and attitude should be positive and between barriers and self-efficacy should be negative. The results obtained in the study confirm these hypotheses. The one exception was a weak non-significant negative correlation between severity and threat which we expected it to be positive. Additionally, we performed bivariate correlation analysis between the reported stage of change and days of physical activity. This correlation was significant ($r = .41$, $P<0.05$).

8.1.1.3 Discussion

The purpose of the study was to test the causal relationships between the constructs in the COMBI model of behavior change. The results demonstrated strong significant correlations between susceptibility and threat, and between the number of days of physical activity and threat, attitude, self-efficacy, mood and cues. The implication of this finding is that the constructs at the 2nd level (that directly influence awareness, motivation and commitment) should be given a high priority, as they seem crucial to the process of behavior change. For example, if the value of attitude is also low and the value of emotions contributing to attitude is low, an intervention directly targeting attitude should have a higher priority. Furthermore, the strong positive correlation between stage of change and days of physical activity indicates that the questions addressing stage of change is a good measure of physical activity behavior.

It is important to note that the participants of the present study were healthy subjects while in the designed model for chronic patients the construct ‘cues’ includes pain and inconvenience resulting from disease, which may result in a different influence and operationalisation compared to the physical cues experienced by healthy subjects. The inclusion of healthy patients can be considered one of the limitations of the present study. Although the model may be applied to healthy subjects, it was developed specifically for chronic patients. Therefore, some minor adjustments may be needed in order to apply it to healthy subjects. Another limitation of the study is that the sample was not representative of the whole population as the questionnaire was spread foremost among the representatives of the highly educated group of university staff. Furthermore, the majority of the constructs were measured with the help of one or two questions, which is arguably not sufficient to grasp the real ‘value’ of a construct for each participant.
In conclusion, this study offers some empirical support for the COMBI behavior change model. The results generally, but not totally, support the way the COMBI model was constructed based on psychological literature.

8.1.2 Study 2

This study was conducted in support of the validation study discussed in Subsection 8.1.1 and was conducted with chronic patients. The data for the present analysis was obtained from the volunteers participating in the pilot study designed for the evaluation of the eMate coaching system that incorporates the COMBI model. Due to the fact that none of the participants opted for coaching on medication, the data were restricted to physical exercise and food intake domains.

8.1.2.1 Method

A total of 17 chronic patients participated in the study by filling out an anonymous questionnaire online. The questionnaire consisted of 92 questions and targeted therapy adherence with respect to two domains: physical exercises and healthy diet. The questionnaire was administered per email and was part of the pilot study for testing the coaching system as a whole (see Section 8.2). It was a modified version of one applied in the previous study. Again, participants were assured of their anonymity and the voluntary character of the survey. No incentives for participation were offered. The study involved 14 males and 3 females. Age ranged between 28 and 80 (Mean = 51.8, SD = 13.8). There were 10 patients with cardiovascular disease and 7 patients with diabetes type 2. The Body Mass Index of the participants ranged between 19 and 37. Of the 17 participants 7 were overweight and 3 obese. The entire questionnaire took approximately 30 minutes to complete.

All of the measures were in Dutch. The questionnaire measured physical activity and healthy diet behaviors and their possible determinants. The 16 determinants of behavior as described by the COMBI model were assessed for two domains, as well as the stage of behavior change per domain (precontemplation, contemplation, preparation, action and maintenance). Besides the UPCC, SQUASH and PANAS scales used in the original questionnaire, two additional externally validated questionnaires were incorporated: PASE and FFQ (see Table 8.2). The other determinants were measured in the same way as in Study 1.

The hypothesis was that we would find positive correlations between all determinants that are connected by causal relations, except for the connection between barriers and self-efficacy (which was expected to be negative).

8.1.2.2 Results

Several sets of bivariate Pearson Product-Moment correlations were performed between the determinants that are positioned close to each other in the model. Additional analysis focused on the relations between the participant’s behavior score per domain (obtained from PASE and FFQ questionnaires), stage of change, and Body Mass Index (BMI). For the related constructs in the model and the additional analysis Pearson product-moment correlations were performed. The results per domain are presented below.
Table 8.2: Validated questionnaires used for the intake questionnaire

<table>
<thead>
<tr>
<th>Construct</th>
<th>Questionnaire</th>
<th>Source</th>
<th>Study</th>
</tr>
</thead>
<tbody>
<tr>
<td>skills</td>
<td>UPCC: Utrechtse Proactive Coping Competence list</td>
<td>Bode, Thoolen, and De Ridder (2008)</td>
<td>1,2</td>
</tr>
<tr>
<td>mood</td>
<td>PANAS: Positive and Negative Affect Schedule</td>
<td>Watson, Clark, and Tellegen (1988)</td>
<td>1,2</td>
</tr>
<tr>
<td>self-efficacy</td>
<td>Question adopted from scale to measure perceived self-efficacy in people with arthritis</td>
<td>Lorig, Chastain, Ung, Shoor, and Holman (1989)</td>
<td>1,2</td>
</tr>
<tr>
<td>physical activity</td>
<td>Dutch Short Questionnaire to Assess Health-Enhancing Physical Activity (SQUASH)</td>
<td>(Wagenmakers, Akker-Scheek, Groothoff, Zijlstra, Bulstra, Kootstra, van Raaij, and Stevens, 2008)</td>
<td>1,2</td>
</tr>
<tr>
<td>physical activity</td>
<td>PASE: Physical Activity Scale for the Elderly</td>
<td>Washburn, McAuley, Katula, Michel, and Boileau (1999)</td>
<td>2</td>
</tr>
<tr>
<td>food intake</td>
<td>F&amp;V FFQ: Fruit and Vegetables Food Frequency Questionnaire</td>
<td>Assema, Brug, Ronda, Steenhuis, and Oenema (2002)</td>
<td>2</td>
</tr>
</tbody>
</table>

Physical Activity  Significant positive correlations were found between skills and self-efficacy ($r = .49$, $P<0.01$), physical activity and threat ($r = .65$, $P<0.01$), and physical activity and cues ($r = .6$, $P<0.05$). The only negative correlation was found between coping and physical activity, but it was non-significant. As far as the correlations between the PASE score, stage of change and BMI of the participants are concerned, only the correlation between BMI and stage of change of physical activity was found significant ($r = -.49$, $P<0.05$).

Healthy Food Intake  Significant positive correlations were found between emotions and attitude ($r = .63$, $P<0.01$), and between number of days of fruit and vegetables intake and a) attitude ($r = .65$, $P<0.01$), b) self-efficacy ($r = .67$, $P<0.01$), c) positive mood ($r = .51$, $P<0.05$), and d) cues ($r = .52$, $P<0.05$). Significant correlations were also found between BMI and stage of change for healthy food intake ($r = -.54$, $P<0.05$), and between fruit and vegetables intake and stage of change ($r = .70$, $P<0.05$).

8.1.2.3 Discussion

The sample size of 17 patients is not sufficient to draw definite conclusions about the validity of the COMBI model for chronic patients, but it demonstrates that the general trends in the relations between the model constructs are correct. However, unexpected trends, though very weak and non-significant, were found between barriers and self-efficacy and between physical activity and coping for physical activity domain. After more detailed examination it became clear that the barriers variable was positively skewed and had small variance. Because the majority of the subjects reported only one or two barriers, this fact may have had consequences for the results. The same holds for the coping variable, which was also found to be not normally distributed and is positively skewed.

For the healthy diet domain, only one unexpected trend was observed. It concerns the relation between negative mood and days of fruit and vegetables intake. This correlation was found positive, though non-significant. Another consistent finding is the correlation between BMI and the stage of behavior change that was found significant both for physical activity and healthy food intake domains. This finding is in line with the general observation reported in the literature that BMI is one of the important indicators of physical activity and diet behaviors and thus should be
8.2.2 Agent-based support for behavior change

included in the analysis of behavior change. The absence of a significant correlation between the PASE score and BMI is also found in a validation study of the PASE questionnaire (Washburn et al., 1999). Washburn, Smith, Jette, and Janney found that the PASE score did not have any significant relationship with body fat percentage, though the PASE score was significantly correlated with the general physical fitness of the subjects (Washburn et al., 1993).

8.2 The eMate System

The COMBI model is incorporated in an intelligent coaching system, called eMate. The eMate system aims to support patients with Diabetes Mellitus type 2, HIV or cardiovascular disease in adhering to their therapy, which can consist of lifestyle advice and/or precise instructions for medication intake. eMate is designed as a ‘cooperative assistant’: an assistant with a coaching character that is able to explain and educate, and expects high participation of the user. Research shows that cooperative assistants might be more effective than ‘direct assistants’, i.e., assistants with an instructing character with brief reports and low expectations on participation (Henkemans, Rogers, Fisk, Neerincx, Lindenberg, and van der Mast, 2007). The eMate system operates as a coach, using both a mobile phone and a website to interact with the user. A mobile phone application supporting both the Android and iPhone platform has been developed that is able to send questions and messages to the user. The interface of the system is designed to suit the needs of usage for older users (a large group among diabetes type II and cardiovascular patients). The interface was designed using insights from usability testing research that elderly users are often easily distracted, have difficulty reading small fonts and closely spaced text, and may have difficulty remembering previously presented information (Barnum, 2011). The design of the app was therefore kept simple but clear, using for instance large fonts and minimizing the need for scrolling. The system performs the following tasks, which are described in the corresponding subsections:

1. determining the stage of change of a user;
2. monitoring the behavior of the user to determine the level of adherence;
3. reasoning about the changes required for improvement;
4. trying to change the user’s perception of specific psychological aspects;
5. updating the beliefs about the user.

8.2.1 Determining stage of change

When someone starts using the system, he or she has to fill out a questionnaire about his or her behavior and beliefs in the domains of medication intake, healthy food intake and physical activity. The questionnaire is based on validated instruments, as listed in Table 8.2. The answers are used to calculate the stage of change of behavior in each of the domains. For each of the constructs, the survey answers that correspond to this construct are converted (using the survey keys) to a single value. These values are stored in the system and taken as starting point for the personalized coaching.

8.2.2 Monitoring level of adherence

In order to provide personalized coaching, it is required to monitor the behavior of the user. The system does this in different ways for the three domains on which it provides coaching. The medication intake behavior of the user is measured directly with the help of an electronic
pillbox. This pillbox registers when it is opened and immediately sends a message via built-in GSM technology to the eMate server, which registers the timestamp of the intake in its database. While the opening of a pill box is no proof of medication being taken, it is a good measure for the level of adherence (Cramer, Mattson, Prevey, Scheyer, and Ouellette, 1989; Deschamps, Geest, Vandamme, Bobbaers, Peetermans, and Wijngaerden, 2008).

The food intake behavior is monitored via a graphical question on the mobile phone (Figure 8.2c). The interface presents 6 icons to the user which represent respectively (left-to-right, top-to-bottom): drinking fluids, eating meat or fish, eating vegetables, eating starchy products, eating fruit and eating snacks. The question that the user has to answer is to what extent he or she behaved in a healthy way with respect eating/drinking from these categories. The three answer options per category are: unhealthy (no click on the icon, icon is not filled with background color), moderately healthy (one click on the icon, icon is half-filled with background color), or healthy (two clicks on the icon, icon is fully filled with background color). Initially, the systems sends the question on a daily basis in the evening hours, but after one week the frequency is adapted to the ‘moving average’ of the answers: users that in a healthy way in the past week only receive the message once in the three days, moderately healthy eaters receive the message once in the two days, and unhealthy eaters receive the question daily. This has been done after initial user-feedback that some of this question was asked too often. Since the feedback is more relevant for users that behave sub-optimal on this aspect, it was decided to reduce the frequency for people that behave well. The website offers the users the possibility to fill in or change their answers about the food intake for the current day or the two days before.

Physical activity is monitored via a combination of a web-based calendar and mobile phone questions. The calendar allows the user to register the number of minutes of physical activities in a weekly schedule for three different categories: 1) walking, 2) cycling, and 3) doing sports. The idea behind these categories is that walking and cycling are (at least in the Netherlands) activities that usually take place on regular basis, e.g., during commuting between home and work. For each entry in the calendar, the user is able to specify whether it is a “confirmed” activity or an activity that still requires confirmation afterwards. For activities in the latter category, users receive a telephone message with the question if he/she actually performed the activity a day after the scheduled physical activity (see Figure 8.2a). When the user did not register any activities on a specific day, the user is asked to recollect how many minutes of activity were performed. The level of adherence is calculated by comparing the average minutes of physical activities with the health goals.

8.2.3 Reasoning about required changes

The aim of the eMate system is to help the user to behave fully compliant with the prescribed therapy. If monitoring reveals that adherence to the therapy on a specific domain (food, medication, activity) is not optimal, the system will use the COMBI model to reason about the underlying causes for non-adherence. It will determine what prevents the user from moving to the next stage of change in that domain by investigating whether the constructs that influence the consecutive stage (that is, the determinants that are related to this stage by an incoming arrow) are a bottleneck. A bottleneck is defined as the construct in the graph that prevents a patient from progressing from one stage of change to another. For example, if a patient is in the preparation phase, the system will investigate the constructs that are related to the action phase. More specifically, it will

1http://www.evalan.com/projects/medication
2The eMate website contains information about what is considered ‘healthy behavior’ with respect to food intake, exercise and medicine intake.
investigate the constructs that are in the paths connected to the action phase. A path consists of a sequence of edges which connect a sequence of constructs. The constructs at the level directly above the stages are called the problem areas. For example, in Figure 8.1 it can be seen that if someone was in stage P, his/her problem area would be commitment. One path connected to this problem area consists of attitude and pros/cons.

The system poses one or two questions via the mobile phone for each of these potential bottlenecks, or hypotheses. Figure 8.2b shows such a question for investigating the patient’s perception on the severity of his/her illness. If the answer shows that this construct might indeed be problematic, the subsequent constructs in the path are investigated. A semi-formal notation of this process is described by Algorithm 1.

This mechanism is an example of model-based diagnosis Davis (1984). By using the approach that is sketched above on a causal model, the system ignores constructs that are not relevant for explaining the current behavior. This has a practical advantage for the user, because only those questions are asked that are necessary to determine the bottleneck that obstructs the healthy behavior. For the purpose of model-based reasoning, the simulation model has been translated into a rule-based representation that allows for backward reasoning over the psychological constructs in the model. To achieve this, the rules connect constructs in the model that have an ‘influence’-relation. That is, if there is an arrow from construct A to construct B in the model (see Figure 8.1), a rule describes that a low value of construct B could be caused by a low value A. Formally, this is done by giving all constructs an attribute ‘is_hypothesis’, which can be set to true or false. The rules specify that a construct is a hypothesis for the bottleneck if its value is below some threshold and the preceding construct on the path is also an hypothesis. The actual values for constructs that are set as hypothesis will have to be investigated via questions posed to the user. This is also specified as a rule. For example, the following two rules connect threat and severity and define that the value for severity should be investigated if it is an hypothesis with an out-of-date value:

\[
\text{IF threat has_value} < 0.5 \& \text{threat is_hypothesis} \Rightarrow \text{TRUE THEN severity is_hypothesis} = \text{TRUE}
\]

\[
\text{IF severity is_hypothesis} \Rightarrow \text{TRUE} \& \text{lifetime} > \text{TIME_TO_LIVE} \Rightarrow \text{THEN severity.investigate}
\]

Using the keys of the validated questionnaires the user’s answers to the questions are translated to values for the construct to which it belongs. These values are stored in a database along

---

**Figure 8.2: Screenshots of the eMate telephone app, original text is in Dutch**
Algorithm 1 Find the hypothesized bottlenecks for behavior change in a domain

\[ C \leftarrow \text{the set of all constructs in the model graph} \]
\[ S \leftarrow \text{the ordered set of all stages of change:}\{PC < C < P < A < M\} \]
\[ s_i \leftarrow \text{the current stage of change of the user, } s_i \in S \]
\[ s_j \leftarrow \text{the stage that directly succeeds } s_i, s_j \in S \]
\[ \tau_i \leftarrow \text{the threshold for construct } i \]
\[ l_i \leftarrow \text{the lifetime for a value of construct } i \]
\[ \text{bottleneck} \leftarrow \text{list of bottlenecks, initially empty} \]

\[ \text{for all } c_k \in C \text{ do} \]
\[ \quad \text{if } \text{connected}(c_k, s_j) \text{ AND age}(c_k) < l_k \text{ then} \]
\[ \quad \quad \text{investigate}(c_k) \]
\[ \quad \text{end if} \]
\[ \text{end for} \]

\[ \text{function investigate}(\text{construct } c_i) \]
\[ \quad \text{update } c_i \]
\[ \quad \text{if value}(c_i) < \tau_i \text{ then} \]
\[ \quad \quad \text{bottleneck} \leftarrow \text{bottleneck } + c_i \]
\[ \quad \text{for all } c_{j \neq i} \in C \text{ do} \]
\[ \quad \quad \text{if connected}(c_j, c_i) \text{ AND age}(c_j) < l_j \text{ then} \]
\[ \quad \quad \quad \text{investigate}(c_j) \]
\[ \quad \text{end if} \]
\[ \text{end for} \]
\[ \text{end if} \]
\[ \text{end function} \]

with a timestamp of their determination. This way, the system maintains an up-to-date representation of the mental state of the user. The reasoning is performed separately for all domains for which the patient has requested coaching. However, the values for mood, cues, skills, severity, susceptibility and threat are considered to be equal across the different domains, and their value is automatically propagated to the other domains via the rules.

The algorithm is implemented in a Java-based rules engine (Drools)\(^1\). The system has 92 rules that specify the reasoning process, which is run on a hourly basis. This results in a set of bottlenecks: the determinants that probably have to change in order for the user to proceed to the next stage of change.

8.2.4 Updating the beliefs about the user

The eMate system provides continuous support to the user. This implies that it should regularly update its belief about the user. This is done in two different ways. First, when the measured behavior is different than the derived stage of change from the user’s answers to the questions (e.g., the user answered questions that indicate he/she behaved in a healthy way yet the monitoring reveals different), the stage of change is set to the more appropriate state. Specifically, if the user is behaving unhealthy, he is assumed to be in the PC, C or P phase, which is further narrowed down via questions. Similarly, when the user is behaving healthy, he is assumed to be in the A or M phase. The second update mechanism works with a specified lifetime that all constructs have assigned to them. As some constructs are more dynamic than others, this values differs per construct. When the lifetime has expired, the system will not use the value from the database, but – when needed in the reasoning about hypotheses – redetermine its value via questions to the user.

\(^1\)http://www.jboss.org/drools/
8.3 Agent-based support for behavior change

8.2.5 Changing the user's perception

The aim of the eMate system is to support and advise the patient in a personalized manner. The intervention therefore targets the bottlenecks as determined by the reasoning process. That is, the user will receive motivational and informative messages related to the problematic constructs. This is done on a weekly basis, together with feedback based on the results of the monitoring. The user thus gets an update of the relation between his goals and his actual behavior, in addition to personalized persuasive messages. An example is provided in Figure 8.2d.

In case multiple factors should be targeted, the system prioritizes the messages in order to support the user most effectively. Once the bottlenecks are established, eMate prioritizes them according to their ‘urgency’ (i.e. how low the value of the construct is for the user) and their ‘changeability’ (to what extent the user is able to change this construct). For example, the perceived social norms are more difficult to change than knowledge of pros and cons of adopting a new behavior.

Due to the model-based reasoning, eMate is able to address the right problems at the right time. However, in order to persuade a user, the formulation of the queries and messages are important as well. All messages are designed in such a way that the user won’t be annoyed or bored by lengthy information messages (this approach is typical for tailored health messages that are commonly used in web-based solutions (Kreuter, Farrell, Olevitch, and Brennan, 2000)). Specifically, eMate sends messages that can be read in its entirety on the display of an average smartphone (3.3 inch), thus minimizing the user’s need to scroll. Furthermore, eMate automatically composes the messages from three separate components: a status update for the user; a motivational message targeting a specific bottleneck (the person’s bottleneck according to the reasoning process); and a link to the relevant part of the website for more information. For each of the components, the systems contains around three templates per domain (i.e. medication / physical activity / food intake) and disease. The template allows for personalization because the name of the person can be filled in and it can be formulated in a formal way (targeting older people) or informal way (better suited for younger people). In total the system contains 264 message components. Each time a message should be generated, the system randomly chooses between components that haven’t been used for this user (or that are sent the longest time ago if all have been sent). Because of this, the user will rarely receive the same message twice. Furthermore, the motivational messages adhere to the principles of motivational interviewing, which have proven to be effective for purposes of coaching and therapy (Miller and Rollnick, 1991). These principles focus on the social functioning of the user and on providing feedback by giving advice and direction. Expressing empathy, cheering and complimenting, and the support of self-efficacy and optimism, are some examples of the principles that are incorporated by the eMate system.

On the website an overview is given of the extent to which the user has reached his/her goals in the past week, which is represented as a percentage and an iconic thumb. See Figure 8.3a for an example. The website also displays the user’s performance on the three different domains. For example, users can see how well they are doing with regard to their medicine intake, as shown by Figure 8.3b, or which activities they have scheduled. Finally, both the unanswered questions and new messages are shown, and it is possible to look up older messages and questions.

---

1 The system was designed for a use of approximately three months, which was judged long enough to establish a behavior change (at least to some degree) with respect to lifestyle.
8.3 System Evaluation

This section describes a preliminary evaluation of the eMate system, using the data obtained in the second study as described in Section 8.1.2. The setup of this (feasibility) study is not suited for drawing conclusions about the uptake and effect of the system. Some participants encountered technical issues related to the specific version of their smartphone, which forced them to stop half-way, and changes in the system have been made during the intervention to solve urgent problems. Therefore, the current evaluation focuses on the functioning of the reasoning system. An evaluative study of the effectiveness of the coaching (in terms of faster or more permanent behavior change) and user experience is currently being conducted.

As described in Section 8.1.2, the study includes data of 17 participants with a chronic illness. The participants were recruited from the general public via an website; a press release has been sent by the university to draw attention to this website. Some 40 people have created an eMate account in the first two months after the public release: about half of them downloaded and installed the app, and 17 of them also filled out the online questionnaire. Filling out the questionnaire was required in order to use eMate, but judging from the dropout of 3 people it was maybe considered a barrier for using the system. The preliminary evaluation is performed by comparing the conclusions of the reasoning process about the bottlenecks with the answers to the questions on physical exercising and food intake that the participants provided via the survey. As stated previously, none of the participants opted for coaching on medication and thus the electronic pillbox was not used in this study.

8.3.1 Analysis of the reasoning process

The essence of the eMate reasoning process is that it identifies the problem areas and bottlenecks with respect to the behavior of the user. Several observations can be made about the reasoning process.

**Trend 1** eMate identified commitment as the problem area for behavior change for most people.

Commitment was found to be the problem area for most people, namely in 65% of combined analyses of exercise and food intake. Motivation was identified as problem area in 29% of the cases, whereas awareness was deemed problematic in only 6%. Table 8.3 shows the bottlenecks in their path for each identified problem area.

**Trend 2** eMate identified coping, social norms and cues most often as the bottlenecks for commitment, motivation, and awareness, respectively.
Table 8.3: Overview of the number of times constructs were identified as bottleneck in their path given a specific problem area.

<table>
<thead>
<tr>
<th>COMMITMENT</th>
<th>MOTIVATION</th>
<th>AWARENESS</th>
</tr>
</thead>
<tbody>
<tr>
<td>coping</td>
<td>social norms</td>
<td>cues</td>
</tr>
<tr>
<td>social norms</td>
<td>self-efficacy</td>
<td>threat</td>
</tr>
<tr>
<td>mood</td>
<td>threat</td>
<td>susceptibility</td>
</tr>
<tr>
<td>emotions</td>
<td>susceptibility</td>
<td>severity</td>
</tr>
<tr>
<td>pros_cons</td>
<td>severity</td>
<td></td>
</tr>
<tr>
<td>attitude</td>
<td>emotions</td>
<td></td>
</tr>
<tr>
<td>skills</td>
<td>attitude</td>
<td></td>
</tr>
<tr>
<td>barriers</td>
<td>pros_cons</td>
<td></td>
</tr>
</tbody>
</table>

For 64% of those who had low commitment values (i.e., a value < 0.5 as calculated by eMate), the construct coping strategies was the hypothesized bottleneck found in the connected paths. In 40% of the cases where motivation was low, eMate identified social norms as the bottleneck. In case of low awareness, the main (and only) bottleneck was cues, i.e. the absence thereof (100%). In 74% of the cases, the identified bottleneck turned out to be the construct with the lowest value in all paths connected to the problem area. However, only for 41% was the identified bottleneck the construct with the lowest value in the entire graph. More specifically, although cues and coping are often the lowest constructs in the graph (44% and 50%, respectively), only coping was often identified as a bottleneck: in 59% of the cases (vs. in 7% of the cases for cues). 12 out of 17 people (71%) had the same underlying problem area for exercise and food intake.

**Trend 3** According to self-reports on healthy behavior most people were in the maintenance phase, both with respect to regular physical exercise and healthy food intake.

From their answers to the questions about their stages of change, it can be derived that 35% of all participants were in the maintenance stage with respect to regular physical exercise. Furthermore, 47% was in that same stage with respect to healthy food intake. This suggests that (i) people are (unrealistically) optimistic about their own adherence, and (ii) therapy adherence is for many people not domain-specific.

From these results several more general observations can be made. First, eMate’s hypotheses of the bottleneck often correspond with low (updated) values of the bottleneck construct. This means that the constructs that end up as targets of interventions are often evaluated by the participant as one of the most problematic determinants for behavior change.

Second, in the cases where the hypothesis turned out not to be the construct with the lowest value in the causal path connected to the problem area, eMate arguably finds a better candidate for intervention, as the search algorithm prunes out constructs that are not relevant. To illustrate this, let us take a closer look at the model of behavior change for Alan (see Figure 8.4). The figure shows for all constructs whether they, given the most recent information, were found to be below or above the threshold. eMate has deducted that Alan is in the contemplation stage and that motivation is his underlying problem area, as the value of this construct is too low for Alan to make the transition to the preparation stage. The connecting paths contain several constructs with low values: severity, emotions and attitude. In the case of Alan, eMate first hypothesizes that attitude is the bottleneck, as it is the only influencing node with a value that was below the threshold. This hypothesis will be investigated. In case the resulting value is above the threshold, eMate will start looking for bottlenecks in other connected paths. If attitude is indeed below the
threshold, eMate will look at the originating nodes for a more detailed hypothesis. In the current scenario, the old value of *emotions* was below the threshold and therefore is the new hypothesis. Suppose that in case of Alan it turns out *emotions* is indeed below the threshold, but is higher than (the outdated value of) *severity*. Considering that Alan’s *threat* has no problematic value, his stage is not likely to improve with the help of interventions targeted at *severity*, even though its value could be lower than that of *attitude* or *emotions*.

Third, eMate does not always identify bottlenecks that are at the end of the paths. Again, this is because eMate prunes the graph such that it targets constructs that are most likely to establish a transition to the next behavior change stage. For example, Alan could have a very low value for *cues*, but since his main problem area is *motivation*, the path containing *cues* will not be explored. By targeting *attitude*, eMate maximizes the chance of increasing Alan’s motivation and helping him reach the next stage.

### 8.3.2 Comparing reasoning conclusions with survey answers

A second type of evaluation of the reasoning performed by eMate was done by using some of the answers to the survey as input for a simulation of the reasoning process, and comparing the outcome of this process with other survey answers. This process consisted of several steps. First, the answers of the 17 participants on questions related to the leaf nodes of the model (i.e. the constructs at the top level inside the dashed-box in Figure 9.1) were used to calculate numerical values for these constructs, normalized to a scale between 0 and 10. Second, values for the constructs at the second level of the model (e.g., threat, attitude) were calculated. However, this calculation did not use the answers to the survey, but used an algebraic combination of the values of the constructs in the first row that are on its path. The values for third level constructs (e.g., awareness, commitment) were calculated in a similar way by combining the values of the constructs on their causal path. Eventually, the calculated values for the second and third level
Agent-based support for behavior change

Table 8.4: Average errors of derived construct values compared with survey answers.

<table>
<thead>
<tr>
<th>method</th>
<th>2nd level</th>
<th>3rd level</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 (average)</td>
<td>2.23</td>
<td>4.64</td>
</tr>
<tr>
<td>2 (maximum)</td>
<td>1.80</td>
<td>1.68</td>
</tr>
<tr>
<td>3 (minimum)</td>
<td>3.79</td>
<td>7.89</td>
</tr>
<tr>
<td>4 (weighted)</td>
<td>1.77</td>
<td>3.15</td>
</tr>
</tbody>
</table>

constructs were compared with the values derived from the survey. This resulted in an error for each construct (defined as the absolute difference between the calculated value and the derived value) for each user and an average error per level.

This simulation was performed using four different algebraic combination functions to calculate the values of constructs from the values of constructs on the causal path:

1. the calculated value is a weighted average (with equal weights) of the values of constructs on the causal path;
2. the calculated value is the maximum of the values of constructs on the causal path;
3. the calculated value is the minimum of the values of constructs on the causal path;
4. the calculated value is a weighted average of the values of constructs on the causal path, where the weights of constructs on the causal path that have a strong impact are larger than of those with a weak impact.1

Table 8.4 presents the different average errors. A few observations about the errors can be made. First, one can see that usually the error at the second level is smaller than the error at the third level. This is as expected because third level calculations are based on the second level calculations which already include some errors. Second, one can see that method 3 (minimum value) results in the highest errors. This observation can be explained by seeing that method 3 is based on the assumption that a good score (= low value) for just one construct on the causal path is sufficient to result in a good score on the resulting construct. However, this assumption is usually not met, since the causal path seems to implicate that high values of other constructs on the path are conditional for a high construct value. Finally, it can be observed that method 2 (maximum value) results in the lowest error (< 1.8 on a scale from 0 to 10). This is in line with the expectation that constructs on the causal path form necessary conditions that need to have a value above a certain threshold.

Overall, these simulations show that the model is able to relate values of leaf constructs to values of constructs at deeper levels. A reasoning process could use the structure of the model and the interpretation of ‘necessary conditions’ for the relation between a construct and the constructs on its causal path.

8.3.3 Limitations

The analyses described above demonstrate that the model can be used to reason about the user’s problem domain and the underpinning theories related to it. The system can accordingly maximize the likelihood that interventions are targeted where they are most needed. We are however well aware that both evaluations only provide initial evidence for the correctness of the reasoning process; the number of participants is not sufficient to extract any significant findings. For a

---

1Specifically, for each participant the survey construct values on the causal path $A_1, \ldots, A_n$ were compared with the survey construct values under investigation $B$. The construct $A_i$, with the value closest to the value of construct $B$ received a large weight ($1 - 0.1 \times n$), all other constructs received a weight of 0.1.
conclusive evaluation, further experiments, as well as an elaborate user interaction analysis, are needed to review eMate’s functioning and effects.

### 8.4 Discussion & Conclusion

This work presents the design, evaluation and use of an autonomous coaching system that supports behavior change for chronic patients. Chronic illness is becoming a worldwide threat to human health. In order to prevent the onset of chronic diseases and to minimize the serious complications that come along with them, an (pro)active approach is in order. The system is called eMate and uses a mobile phone application, a website, and an electronic pillbox to interact with the user and to determine his or her current stage of behavior change. COMBI, an integrated cognitive model for behavior change, forms the core of eMate and represents the dynamics of internal processes underlying behavior change. COMBI is an example of a causal modeling approach to develop complex, user-tailored interventions. eMate differs from existing approaches in that it targets the user’s motivation and interests, and tailors intervention messages on the basis of the underlying mechanisms of behavior change. Model validation studies demonstrate that the core of the embedded cognitive model COMBI corresponds to reality, and preliminary results show that the model can be used for intelligent reasoning and tailored interventions initiated by the system. In the future, the model could also be used to predict the effect of an intervention, in order to let the system choose the most effective one.

Although interface design was not the point of focus of this study, system design and user experience is extremely important for the successful use of computer support systems. In order to match the design of the eMate interface to the target audience, prototypes of both the app and the website have been discussed with professionals that have experience with the use and design of apps for both chronic patients and elderly people. There were repeated design rounds and formative testing was done while the system was in development. An experiment in which the eMate system is being evaluated with regard to its design and credibility is currently being performed, and a more in-depth study with a large group of participants is being set up. However, the current interface can be much enhanced and be made more attractive for usage. Future work should focus more on this aspect of system design and on its influence on how well users adopt the system.

Although eMate has been developed for chronic patients, it can be used by everyone who wants to change some aspect of his or her behavior since it is easily adjustable and straightforward to use. Calculating the construct values can be adapted for other domains by using surveys appropriate for that domain. The rules and tailored messages can be adjusted to include different conditions and requirements. The general mechanisms for behavior change are similar across a variety of domains, which enables eMate to support for example sunscreen use or sustainable energy consumption. Currently, a study is being conducted in which the eMate system is used to stimulate students to take the stairs more often (instead of opting for the elevator).

Moreover, eMate is not only a helpful coach for people to improve their health, but also a resourceful tool for researchers who want to test their theories of behavior change and persuasive interventions. Currently, in the field of eHealth the flexibility and adaptivity of models and frameworks to different domains is still very limited. This, however, is an area where important gains can be achieved. Developing these models and frameworks takes great time and effort, and by designing them in a way that is not restricted to a single domain will highly improve their usability.
In the present work, the user’s social environment is only taken into account to a limited extent. It is obvious that systems that make use of more detailed knowledge of the relation between the social environment and behavior have a large potential. This can be facilitated by using information from online social networks. Also, including components such as GPS, physical movements sensors or stress and mood detectors will improve its functionality and increase the possibilities for tailored messages and context-specific interventions.

The main limitation of the present work is that it presents only preliminary results concerning the validation of the underlying model of behavior change and the evaluation of the system that incorporates this model. Follow-up studies are currently being conducted with groups of patients in an experimental setup that cover these limitations. Using the data from these experiments, the weights and thresholds used to reason with the model (as elaborated in Section 8.2) can be initialized in a more meaningful way. For example, different user profiles could be distinguished for which different constructs might be important. Additionally, the effectiveness and user preference of how he/she is addressed by the system (i.e., the persuasive strategies) could be learned based on interactions with the user.

Acknowledgements

This work is supported by the Netherlands Organization for Health Research and Development (ZonMW) programme “chronic disease management”, grant number 300020005. Thanks are due to our partners in this project, as well as to Ingeborg Brouwer, Anne Heideveld, Bart Kamphorst, Annemiek Linn and Marjolein Visser for valuable comments.

References


