Abstract
The adoption and maintenance of a healthy lifestyle is a fundamental pillar in the quest towards a healthy society. Modern (mobile) technology allows for increasingly intelligent systems that can help to optimize people’s health outcomes. One of the possible directions in such mHealth systems is the use of intelligent reasoning engines based on dynamic computational models of behavior change. In this work, we investigate the accuracy of such a model to simulate changes in physical activity levels over a period of two to twelve weeks. The predictions of the model are compared to empirical physical activity data of 108 participants. The results reveal that the model’s predictions show a moderate to strong correlation with the actual data, and it performs substantially better than a simple alternative model. Even though the implications of these findings depend strongly on the application at hand, we show that it is possible to use a computational model to predict changes in behavior. This is an important finding for developers of mHealth systems, as it confirms the relevance of model-based reasoning in such health interventions.
This chapter appeared as:

6.1 Introduction

It is well known that engaging in sufficient physical activity has many beneficial effects on physical and mental health (Conn et al., 2011; Eime et al., 2013). On the contrary, low levels of physical activity have been associated with increased risks of cardiovascular diseases, cancer, diabetes, and mental illness (Lee et al., 2012). Despite these prominent advantages, a large proportion of the Western population does not meet the guidelines of being moderately to vigorously active for at least 30 minutes on at least five days a week (Haskell et al., 2007).

It is believed that mobile technology provides an opportunity to support people with increasing their level of physical activity (Knight et al., 2015; Payne et al., 2015; Sieverdes et al., 2013), and there is also some initial evidence that this is effective (Stephens and Allen, 2013). However, just monitoring physical activity is not sufficient to achieve a durable improvement (Gierisch et al., 2015). There is a need for development of evidence-informed mobile apps that apply advanced technological features in order to yield long-term effects (Knight et al., 2015).

In (Klein et al., 2015), we described how we developed an intelligent system to stimulate physical activity for young adults. Part of the intelligence lies in the fact that it uses a computational model of behavior change to predict the effect of different intervention strategies on the activity level of the users. The model consists of temporal-dynamic relations between determinants of behavior change. The model predictions are used for deciding on the support messages for specific users in each phase of the intervention. We believe that this can result in a highly tailored and personalized intervention.

The model is based on a number of different theories on behavior change (Michie and Johnston, 2012). Most of these theories have been validated independently, and there also have been validations of integrated frameworks (Cane et al., 2012). However, these validations usually look at correlations between the different constructs in the theories. Hence, we do not yet know whether the dynamic computational representation of our integrated model provides a valid description of the process of behavior change. Therefore, we would like to know to what extent the model is a valid way to predict the most effective coaching strategies. This paper describes a first step towards such a validation. We use empirical data collected in the effectiveness study to compare the prediction of the model based on the initial questionnaire with the actual change in physical activity that has been measured in the study. In addition, we compare the actually measured behavior with predictions of an alternative simple model. The results provide an initial answer about the validity of the model.

The remainder of this paper is organized as follows. In Section 6.2, the details of the computational model under consideration are presented. Section 6.3 describes the methods used to provide a validation of the computational model. The results are presented in Section 6.4, and reflected upon in Section 6.5.

6.2 Computational model of psychosocial influences on physical activity based on social cognitive theory

The computational model investigated in this paper was designed in context of the development of an intelligent behavior change support system (Klein et al., 2015). The reasoning engine of this system uses the model to predict what available coaching strategies are the
most promising to improve the user’s behavior. (For more information, see (Klein et al., 2015).)

The model captures the dynamics between psychosocial influences on physical activity behavior. It describes the relations between several psychological determinants, such as self-efficacy, intentions and social norms, and their influence on physical activity behavior. The model under investigation in this paper is an adaptation of the computational model presented in (Mollee and van der Wal, 2013). Therefore, the concepts and the relations of the model described below are explained in more detail in the original publication. The revision was motivated by a decrease in conceptual detail and computational complexity of the model, and by suggestions of experts in behavior change.

The computational model is largely based on the social cognitive theory by Albert Bandura 1998. This is a well-established theory of behavior change, with high applicability in the domain of health behavior (Bandura, 2004). The theory has proven to account for a large proportion of the variance in physical activity (Rovniak et al., 2002), and is therefore very suitable as a basis to describe the dynamics underlying physical activity behavior.

All concepts are modeled numerically, as real values in the interval [0,1], and the relations are formalized as differential equations. The relations express the influence of the source concept on the target concept by increasing or decreasing the value of the target concept in the direction of the source target, moderated by a parameter named $\beta_{\text{source, target}}$ (or occasionally $\beta_{\text{source1+source2,target}}$). The increase or decrease of the target concept is also relative to its current value: e.g., $SE(t)$ in case of a decrease and $(1 - SE(t))$ in case of an increase. This consideration of the current value also ensures that the concept values stay in the interval [0,1]. The constant $\Delta t$ indicates the step size of the model, and is set at 0.1 to ensure smooth results. Figure 6.1 shows a graphical representation of the dynamic relations between all concepts in the model.

Below, the meaning of the concepts in the model are explained and the formal relations are specified.

**Self-Efficacy:** The self-efficacy (SE) is a key element of the process described by the social cognitive theory. It represents the confidence in one’s own ability to achieve certain goals, which plays a fundamental role in the acquisition and maintenance of some desired behavior. The self-efficacy increases with high satisfaction of the current behavior, and it decreases if one is dissatisfied with his/her behavior.

\[
\begin{align*}
\text{if}(SE(t) \geq Sat(t)) : & \quad SE(t+1) = SE(t) + \beta_{Sat,SE} \cdot (Sat(t) - SE(t)) \cdot \Delta t \cdot SE(t) \\
\text{if}(SE(t) < Sat(t)) : & \quad SE(t+1) = SE(t) + \beta_{Sat,SE} \cdot (Sat(t) - SE(t)) \cdot \Delta t \cdot (1 - SE(t))
\end{align*}
\]

**Impediments:** Impediments (Imp) are the (personal, situational or systemic) factors that form an obstacle to the desired behavior. The self-efficacy plays a role in how insurmountable one views those obstacles. Therefore, the personal impediments (Input_Imp) are adjusted based on the level of the self-efficacy.

\[
\begin{align*}
\text{if}(SE(t) \geq \text{Input}_\text{Imp}(t)) : & \quad Imp(t) = \text{Input}_\text{Imp}(t) - (\beta_{SE,Imp} \cdot (SE(t) - \text{Input}_\text{Imp}(t))) \cdot \Delta t \cdot \text{Input}_\text{Imp}(t)
\end{align*}
\]
6.2 Computational model of psychosocial influences on physical activity

Figure 6.1: Graphical representation of the model.

\[ \text{if}(SE(t) < \text{Input}_\text{Imp}(t)) : \]
\[ \text{Imp}(t) = \text{Input}_\text{Imp}(t) - (\beta_{SE,\text{Imp}} \cdot (SE(t) - \text{Input}_\text{Imp}(t))) \cdot \Delta t \cdot (1 - \text{Input}_\text{Imp}(t)) \]

In the system, these personal ‘input impediments’ are assessed through a questionnaire, so they reflect the user’s overall experience of barriers on a scale of 0 (no impediments) to 1 (very strong impediments).

**Social Norm:** The social norm (SN) represents the behavioral standards that one’s social connections impose on him or her. It is derived directly from information about the user’s social network, and it assumed to be stable for the duration covered by the simulations.

**Long-Term Goals:** The long-term goals (LTG) can be interpreted as the overall motivation to achieve change in the behavior. The levels of self-efficacy can increase or decrease the long-term goals.

\[ \text{Change}_{\text{LTG}}(t) = (\beta_{SE,\text{LTG}} \cdot (SE(t) - \text{LTG}(t))) \]

\[ \text{if}(\text{Change}_{\text{LTG}}(t) \geq 0) : \]
\[ \text{LTG}(t + 1) = \text{LTG}(t) + \text{Change}_{\text{LTG}}(t) \cdot \Delta t \cdot (1 - \text{LTG}(t)) \]

\[ \text{if}(\text{Change}_{\text{LTG}}(t) < 0) : \]
\[ \text{LTG}(t + 1) = \text{LTG}(t) + \text{Change}_{\text{LTG}}(t) \cdot \Delta t \times \text{LTG}(t) \]
**Intentions:** The intentions (Int) denote the user’s aims for the desired behavior. They provide focus and a measure for evaluation. The intentions are influenced by the self-efficacy, the social norm and the outcome expectations, and adjusted by the perceived impediments.

\[
\text{Change}_{-\text{Int}}(t) = (\beta_{SE,\text{Int}} \cdot (SE(t) - \text{Int}(t)) + \beta_{LTG,\text{Int}} \cdot (LTG(t) - \text{Int}(t)) + \beta_{OE,\text{Int}} \cdot (OE(t) - \text{Int}(t)) \\
+ \beta_{SN+Sat,\text{Int}} \cdot (Sat(t) - SN(t)) - \beta_{Imp,\text{Int}} \cdot Imp(t))
\]

if \((\text{Change}_{-\text{Int}}(t) \geq 0)\):

\[
\text{Int}(t + 1) = \text{Int}(t) + (\text{Change}_{-\text{Int}}(t)) \cdot \Delta t \cdot (1 - \text{Int}(t))
\]

if \((\text{Change}_{-\text{Int}}(t) < 0)\):

\[
\text{Int}(t + 1) = \text{Int}(t) + (\text{Change}_{-\text{Int}}(t)) \cdot \Delta t \cdot \text{Int}(t)
\]

**Behavior:** The behavior (Beh) describes the level of physical activity that someone is engaged in: its value is 0 if someone is not physically active at all, and 1 if someone is maximally active. It is mainly influenced by the self-efficacy, outcome expectations, intentions and the impediments.

\[
\text{Change}_{-\text{Beh}}(t) = (\beta_{SE,\text{Beh}} \cdot (SE(t) - \text{Beh}(t)) + \beta_{Int,\text{Beh}} \cdot (\text{Int}(t) - \text{Beh}(t)) \\
+ \beta_{OE,\text{Beh}} \cdot (OE(t) - \text{Beh}(t)) - \beta_{Imp,\text{Beh}} \cdot Imp(t))
\]

if \((\text{Change}_{-\text{Beh}}(t) \geq 0)\):

\[
\text{Beh}(t + 1) = \text{Beh}(t) + (\text{Change}_{-\text{Beh}}(t)) \cdot \Delta t \cdot (1 - \text{Beh}(t))
\]

if \((\text{Change}_{-\text{Beh}}(t) < 0)\):

\[
\text{Beh}(t + 1) = \text{Beh}(t) + (\text{Change}_{-\text{Beh}}(t)) \cdot \Delta t \cdot \text{Beh}(t)
\]

**Satisfaction:** The satisfaction (Sat) denotes one’s perception of his/her own behavior, i.e. an evaluation of the behavior. It is based on the difference between one’s intentions and current behavior, and adjusted with the perceived impediments.

\[
\text{Change}_{-\text{Sat}}(t) = (\beta_{Int+Beh,\text{Sat}} \cdot (\text{Beh}(t) - \text{Int}(t)) + \beta_{Imp,\text{Sat}} \cdot Imp(t))
\]

if \((\text{Change}_{-\text{Sat}}(t) \geq 0)\):

\[
\text{Sat}(t + 1) = \text{Sat}(t) + (\text{Change}_{-\text{Sat}}(t)) \cdot \Delta t \cdot (1 - \text{Sat}(t))
\]

if \((\text{Change}_{-\text{Sat}}(t) < 0)\):

\[
\text{Sat}(t + 1) = \text{Sat}(t) + (\text{Change}_{-\text{Sat}}(t)) \cdot \Delta \cdot \text{Sat}(t)
\]

**Outcome Expectations:** The outcome expectations (OE) represent the anticipated results of performing the behavior, on a physical, personal and social level. They are influenced by one’s satisfaction with the current behavior and the self-efficacy.
\[ \text{Change}_OE(t) = (\beta_{\text{Sat},OE} \cdot (\text{Sat}(t) - OE(t)) + \beta_{\text{SE},OE} \cdot (\text{SE}(t) - OE(t))) \]

\[ \text{if}(\text{Change}_OE(t) \geq 0) : \]
\[ OE(t + 1) = OE(t) + (\text{Change}_OE(t)) \cdot \Delta t \cdot (1 - OE(t)) \]

\[ \text{if}(\text{Change}_OE(t) < 0) : \]
\[ OE(t + 1) = OE(t) + (\text{Change}_OE(t)) \cdot \Delta t \cdot OE(t) \]

The values of all parameters (\( \beta \)) can be adjusted by the modeler. In the current implementation of the model, the parameters were chosen based on correlations between the concepts found in literature (Plotnikoff, Costigan, et al., 2013; Plotnikoff, Lippke, et al., 2008; Rovniak et al., 2002), in order to keep the ratio between the parameters in accordance with empirical findings. This parameter set is shown in Table 6.1. Additionally, one day was chosen to correspond with 10 time steps.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \beta_{\text{Sat,SE}} )</td>
<td>0.50</td>
</tr>
<tr>
<td>( \beta_{\text{Imp,Int}} )</td>
<td>0.08</td>
</tr>
<tr>
<td>( \beta_{\text{SE,Int}} )</td>
<td>1.00</td>
</tr>
<tr>
<td>( \beta_{\text{SE,Imp}} )</td>
<td>0.43</td>
</tr>
<tr>
<td>( \beta_{\text{Imp,Sat}} )</td>
<td>0.25</td>
</tr>
<tr>
<td>( \beta_{\text{Int+Beh,Sat}} )</td>
<td>0.50</td>
</tr>
<tr>
<td>( \beta_{\text{SE,Beh}} )</td>
<td>0.17</td>
</tr>
<tr>
<td>( \beta_{\text{Int,Beh}} )</td>
<td>0.60</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \beta_{\text{Imp,Beh}} )</td>
<td>0.25</td>
</tr>
<tr>
<td>( \beta_{\text{SE,LTG}} )</td>
<td>0.05</td>
</tr>
<tr>
<td>( \beta_{\text{LTG,Int}} )</td>
<td>0.20</td>
</tr>
<tr>
<td>( \beta_{\text{Sat,OE}} )</td>
<td>0.10</td>
</tr>
<tr>
<td>( \beta_{\text{SE,OE}} )</td>
<td>0.05</td>
</tr>
<tr>
<td>( \beta_{\text{OE,Int}} )</td>
<td>0.02</td>
</tr>
<tr>
<td>( \beta_{\text{OE,Beh}} )</td>
<td>0.01</td>
</tr>
<tr>
<td>( \beta_{\text{SN+Sat,Int}} )</td>
<td>0.02</td>
</tr>
</tbody>
</table>

6.3 Methods

In order to assess the validity of the computational model described above, empirical data is collected, preprocessed and analyzed. These steps are described in this section.

6.3.1 Data collection

The data was collected in context of a user study, in which three (versions of) physical activity promotion apps were tested. Each of the participants (\( N = 108 \)) used one of the apps for at least 12 weeks, in the period between March and October 2016.

At the start of the experiment, the participants were asked to fill in an extensive intake questionnaire. This questionnaire included questions about their demographics (e.g., gender, age), about their daily life patterns (e.g., occupation, important locations, travel options), and about psychological concepts underlying behavior (e.g., intentions, self-efficacy). Each of the eight psychological constructs is assessed by a number of items on a four- or five-point Likert scale, or by one item on a scale of 1 to 10. The items were based on extensive, validated questionnaires (e.g., Frank et al., 2009; Rovniak et al., 2002; Sallis et al., 1988).

All participants received a Fitbit One activity tracker that monitored their physical activity and synchronized their data wirelessly to their assigned app. The tracker registers steps, floors climbed, distance, calories burned and active minutes. As mentioned before, the
participants measured their physical activity via the Fitbit One for a minimum of 12 weeks (that is, apart from possible dropouts). The data from the first week was used to assess the initial physical activity, whereas the subsequent eleven weeks were used as ground truth to compare with the model’s predictions.

6.3.2 Data preprocessing

In order to obtain the initial values for the concepts of the computational model, the responses to the questionnaire items for assessing the psychological constructs were aggregated per concept. As the model assumes numerical values between 0 and 1, all responses were rescaled and averaged per concept in order to fit in that same range.

The initial value of the behavior concept was based on the Fitbit step data, rather than self-reported questionnaire answers. This value was calculated in three steps. First, the number of steps in the first seven days of participation was averaged, while discarding any days with no recorded steps (e.g., because the participant forgot to wear the activity tracker). Then, the average number of steps was capped off at 15,000 steps per day, as that represents amply complying with the guideline of 10,000 steps per day. Finally, the initial value of the behavior concept was obtained by normalizing the average number of steps. This way, a behavior value of 1.0 corresponds with 15,000 (or more) daily steps, which is regarded as “optimal” behavior. The calculation of this initial behavior value is shown in Equation 6.1.

\[
\text{Beh}(t_0) = \frac{\min(15,000, \text{NumSteps})}{15,000}
\]

where NumSteps is the average number of steps for the days in the first week, discarding any days with no recorded steps.

6.3.3 Analyses

In order to investigate the validity of the computational model, the assessments of the psychological constructs and the normalized average number of steps in the first week were used to initialize the computational model (as described in Section 6.3.2. Starting from these values, the model was run to predict the physical activity behavior in the second week, third week, etc., up to the twelfth week of the experiment. As for the calculation of the initial behavior value, the actual behavior values for these subsequent weeks were determined as well. Then, the actual change in behavior and the predicted change in behavior were calculated.

The accuracy of the model’s predictions was assessed by comparing the predicted differences to the actual differences in the behavior values. This was done by means of calculating Spearman rank-order correlations.

In order to exclude the possibility that the model’s performance relies on certain underlying pattern in the data (e.g., high values will probably decrease, and vice versa), it was evaluated by comparing it to the results of a simple alternative model. In this ‘random model’, for each user a random value in [0,1] is drawn as predicted behavior value. This random model will also show that high values generally decrease and vice versa, as the probability of drawing a value below (for example) 0.8 is higher than drawing a value above 0.8. Spearman correlations were calculated for the random predictions as well. To avoid
flukes, the random model was applied and evaluated 100 times, and the resulting correlation coefficients and p-values were averaged.

Although 108 participants filled out the intake questionnaire and wore the Fitbit One for an intended period of 12 weeks, not all participants had step data for each subsequent week (e.g., dropouts). Those were not considered in the analyses.

6.4 Results

The 108 people that participated in the user study were between 18 and 30 years old at the time of the data collection. Of those, 22 were male and 86 were female. However, the number of participants with usable data varied each week. Table 6.2 shows the number of users whose data was included in the analyses.

Table 6.2: Number of users included and excluded in the analyses for each predicted week.

<table>
<thead>
<tr>
<th>Week number</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. included users</td>
<td>92</td>
<td>88</td>
<td>87</td>
<td>85</td>
<td>82</td>
<td>78</td>
<td>79</td>
<td>79</td>
<td>74</td>
<td>72</td>
<td>66</td>
</tr>
<tr>
<td>No. discarded users</td>
<td>16</td>
<td>20</td>
<td>21</td>
<td>23</td>
<td>26</td>
<td>30</td>
<td>29</td>
<td>29</td>
<td>34</td>
<td>36</td>
<td>42</td>
</tr>
</tbody>
</table>

The results of the Spearman rank-order correlation tests are summarized in Table 6.3. It shows the correlation coefficient ($r_s$) and corresponding p-value for the predictions of both the computational model and the random model. The results of the random model are based on 100 draws of a random prediction for each user.

Table 6.3: Results of the Spearman rank correlation for week 2 up to week 12.

<table>
<thead>
<tr>
<th>Week</th>
<th>Computational Model</th>
<th>Model (avg)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Spearman’s $r_s$</td>
<td>p-value</td>
</tr>
<tr>
<td>2</td>
<td>.4134</td>
<td>&lt; .001</td>
</tr>
<tr>
<td>3</td>
<td>.4019</td>
<td>&lt; .001</td>
</tr>
<tr>
<td>4</td>
<td>.3001</td>
<td>.0047</td>
</tr>
<tr>
<td>5</td>
<td>.3957</td>
<td>&lt; .001</td>
</tr>
<tr>
<td>6</td>
<td>.3064</td>
<td>.0051</td>
</tr>
<tr>
<td>7</td>
<td>.5522</td>
<td>&lt; .001</td>
</tr>
<tr>
<td>8</td>
<td>.3173</td>
<td>.0044</td>
</tr>
<tr>
<td>9</td>
<td>.3201</td>
<td>.0040</td>
</tr>
<tr>
<td>10</td>
<td>.2841</td>
<td>.0142</td>
</tr>
<tr>
<td>11</td>
<td>.2319</td>
<td>.0499</td>
</tr>
<tr>
<td>12</td>
<td>.3510</td>
<td>.0039</td>
</tr>
</tbody>
</table>

To illustrate, Figure 6.2 shows a scatter plot of the changes in the behavior values of 85 included users as predicted by the model (on the vertical axis) and the corresponding
changes according to the empirical data (on the horizontal axis). The scatter plot for the random model shows the predictions of one of the 100 repeated runs.

![Figure 6.2: Correlation plots with the predictions of the computational model (left) and the random model (right) on the vertical axis and the empirical data on the horizontal axis for week 5.](image)

### 6.5 Discussion

The results presented in Section 6.4 reveal that the computational model performs quite well in predicting the change in physical activity level. The predictions show a weak to moderate positive correlation with the actual data, which is statistically significant ($p < .05$) for all predicted weeks. In contrast, the random model has both weaker and non-significant correlations with the empirical data. However, the random model performs relatively well on some of the weeks (i.e., week 2 and week 7). This indicates that some characteristics of the data can make it easier to predict the change right. Further investigation into this finding should reveal why this happens.

The work presented in this paper clearly is only a first step in the direction of validating computational models that are applied in mHealth systems. For instance, it would be interesting to see whether the computational model is also able to predict the course of the behavior on a more detailed (i.e., daily) level, rather than considering the errors for its predictions per week. On the other hand, the steady good performance for each week suggests that the model captures the dynamics of the behavior over time quite well. Also, the current work is limited to 12 weeks, which might not be transferable to longer periods. The model’s performance on predicting the underlying psychological constructs would be an interesting further exploration as well. Moreover, as Table 6.2 shows, the dataset contained a substantial number of missing observations, for example because of dropouts or participants forgetting to wear the activity tracker. This could have affected the results, and therefore replicating the analyses on a more complete dataset would be another valuable endeavor.

Several directions for further analysis could reveal whether the computational model is able to perform even better than the results found in this work. For example, the current model uses global parameters (based on indications from literature), but it is plausible that better results could be obtained when the parameters are tuned to the users in the dataset,
either globally or individually. In addition, the computational model does not account for the fact that the participants in the study were exposed to an intervention during the data collection: they were using a physical activity promotion app (see Section 6.3.1). By taking a potential effect of the intervention into account, the model’s predictions could arguably be improved even further.

Validation of dynamic computational models is an important endeavor, as it allows researchers to better understand the dynamics of the modeled behavior through simulations. Moreover, this work presents a step in the direction of more reliable and effective mHealth systems. After all, if the computational models underlying their reasoning engines are proven trustworthy, this increases the dependability of the support provided by the mHealth systems.
References


Cane, James, Denise O’Connor, and Susan Michie (2012). “Validation of the theoretical domains framework for use in behaviour change and implementation research”. In: Implementation Science 7.1.


REFERENCES


Plotnikoff, Ronald C., Sonia Lippke, Kerry S. Courneya, Nick Birkett, and Ronald J. Sigal (2008). “Physical Activity and Social Cognitive Theory: A Test in a Population Sample of Adults with Type 1 or Type 2 Diabetes”. In: Applied Psychology 57.4, pages 628–643.


