Part Six

17 Discussion & Conclusion

17.1 Summary of research contributions
17.2 Discussion of research questions and results
17.3 Ethical aspects
17.4 Future work
17.5 Conclusion
The research presented in this thesis explored new approaches to behavior change interventions for physical activity, by applying techniques from modern (mobile) technology and artificial intelligence as well as scientific knowledge from human-directed disciplines within psychology and social sciences. Because of the interdisciplinary nature of this subject, it was addressed by focusing on different aspects from various perspectives.

This chapter reflects on how the previous chapters contribute to the research objective outlined in Chapter 1. In Section 17.1, the overall contributions of the work presented in this thesis are summarized. Section 17.2 presents a discussion of the results, by revisiting the research questions from Section 1.3 of the Introduction. Section 17.3 reflects upon the ethical aspects of the Active2Gether intervention and Section 17.4 looks ahead at future work. Finally, Section 17.5 presents the conclusion of this thesis.

17.1 Summary of research contributions

The research presented in this thesis helps to advance the field of innovative healthy lifestyle promotion, by investigating and developing methods that make coaching more personal, more relevant and overall more ‘intelligent’. Because of the interdisciplinary nature of the research objective, the contributions of this thesis are diverse. This section outlines its overall research contributions, thereby explaining how this research contributes to the advancement of intelligent behavior change applications.

First of all, an inventory is made of the state-of-the-art of apps that promote physically active lifestyles. This is done from two perspectives. First, it is investigated to what extent behavior-change techniques are applied in the current offer of physical activity apps. This gives insight into how evidence-based these apps are. Second, the extent to which technological features are implemented is explored. This helps to understand how much of the technological potential is currently used in such coaching apps. Overall, the research into the state-of-the-art contributes to an understanding of the progress of physical activity coaching apps so far, as well as insight into missed opportunities. In combination with an exploration of user preferences for physical activity coaching apps, this provides a good starting point for further research and development of innovative healthy lifestyle promotion.

Second, some of the research in this thesis contributes to insights in human behavior and factors that influence behavior change processes. This is done, for example, through investigating whether people who are willing to join an online social community in a physical activity promotion program benefit more from it than people who don’t sign up for the community. The results of the analyses do indeed point in that direction. Another new
insight is that it is important to consider people’s preferred direction of social comparison when implementing a social comparison feature, as presenting data of other users for comparison may lead to adverse effects if it contradicts with the user’s preference. Furthermore, computational models of behavior change processes also contribute to understanding human behavior in this area. This thesis presents a computational model of influences on physical activity behavior, which can be a helpful tool to further explore the dynamics through simulations and property analysis. The preliminary validation of this model adds to its reliability, and therefore to the value of its simulations. Finally, two validations of an existing computational model of social contagion help to support the credibility of this model, as well as to understand the role of the social contagion process in collective behavior change attempts.

The work on the design and development of innovative behavior change approaches is the third overall contribution of this thesis. Clearly, the documentation of the overall design and implementation of the Active2Gether system provides valuable information. It may serve as inspiration for future developments of behavior change systems, and the related lessons learnt point to the more or less promising directions. In addition, some behavior change approaches (that may or may not be included in the final version of the Active2Gether system) are explored in more detail. For example, a potential way to apply the computational model of social contagion to achieve behavior change within a social network is investigated. Similarly, secondary aspects of the computational model of influences on physical activity behavior, that come into play when applying it in a real-life behavior change system, are considered. Such explorations help to understand which endeavors might prove helpful in the design of behavior change interventions. Finally, the user evaluation of the Active2Gether system (in comparison with a commercially available app) helps to understand which aspects are appreciated by end users. This information can help to increase engagement, and thereby adherence and overall effectiveness. The fact that the evaluation is based on real use of the apps for a considerable amount of time, rather than based on hypothetical use scenarios, makes it an even more valuable contribution to research on behavior change applications.

The fourth overall contribution worth mentioning is the collection of several datasets in the context of this thesis. These datasets serve different purposes: they form the basis for answering specific research questions, they are used as input to the Active2Gether system that was developed in the context of this thesis, or they are used to evaluate the effectiveness and user appreciation of the Active2Gether system. Overall, the datasets provide very rich information regarding physical activity behavior and a diversity of related factors, such as the users’ significant locations, the availability of stairs at and transport options to these locations, friendship connections between users, and extensive psychological questionnaires about constructs like self-efficacy, intentions, perceived barriers, et cetera. Therefore, they did not only contribute to the research in this thesis, but they allow for a wide range of further analyses that were beyond the scope or time limit of this thesis. These additional investigations will likely lead to further insight in human behavior change dynamics.

17.2 Discussion of research questions and results

The research objective of this thesis is to investigate how mobile technology and artificial intelligence techniques can be applied in the design of a behavior change system that aims
to increase physical activity levels in young adults. In doing so, four subquestions were formulated, and each of the previous parts focused on answering one of these research questions. Table 17.1 shows an overview of the parts, research questions and chapters that together form this thesis. In the following subsections, the research questions are revisited one by one, the results related to these questions are discussed, and the implications and limitations of the work are explored.

Table 17.1: Overview of the parts, research questions and chapters in this thesis.

<table>
<thead>
<tr>
<th>Part I</th>
<th>Part II</th>
<th>Part III</th>
<th>Part IV</th>
<th>Part V</th>
<th>Part VI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Introduction</td>
<td>RQ 1</td>
<td>RQ 2</td>
<td>RQ 3</td>
<td>RQ 4</td>
<td>Discussion</td>
</tr>
<tr>
<td>Chapter 1</td>
<td>Chapter 2</td>
<td>Chapter 5</td>
<td>Chapter 8</td>
<td>Chapter 13</td>
<td>Chapter 17</td>
</tr>
<tr>
<td>Chapter 3</td>
<td>Chapter 6</td>
<td>Chapter 9</td>
<td>Chapter 10</td>
<td>Chapter 14</td>
<td></td>
</tr>
<tr>
<td>Chapter 4</td>
<td>Chapter 7</td>
<td>Chapter 11</td>
<td>Chapter 15</td>
<td>Chapter 16</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Chapter 12</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### 17.2.1 Research question 1

**What are requirements for mobile behavior change interventions for physical activity based on the state of the art of such interventions and user preferences of the target population?**

The research in this thesis started by investigating the state of the art and user preferences for physical activity interventions. This way, a starting point was set for further development of such interventions, and promising directions or possible pitfalls were identified.

Reviewing the actual offer of physical activity apps, rather than the (only sparsely available) literature on this topic, led to valuable insights into the characteristics of currently available apps. The results presented in Chapter 2 show that the analyzed apps on average use five different behavior behavior change techniques, and none of the apps use more than eight or less than two, out of a possible 23 techniques. The most frequently used techniques were providing feedback, self-monitoring and goal setting, whereas some techniques were not used in any of the reviewed apps. Since research shows that the use of behavior change techniques is associated with effectiveness (Webb et al., 2010), these findings suggest that incorporating more of such techniques is a promising endeavor for the development of physical activity apps.

Similarly, the results in Chapter 3 reveal that the reviewed apps on average incorporate eight technological features out of a possible 37, and the numbers ranged between 0.5 and 19.5 (with 0.5 scores indicating disagreement between reviewers). The features that were identified most often were user input (to log activities or to form a personal profile), a textual/numerical overview of the user’s behavior and progress, sharing achievements or workouts in internal or external social networks, and general advice on physical activity. Many features were discovered only rarely, among which adaptation, integration with external sources, feedback based on the user’s physical/social context, encouragement
through gamification or some form of punishment, and the possibility to contact an expert through the app.

When exploring the requirements for mobile healthy lifestyle interventions, the possibilities (in terms of evidence-based behavior change techniques or technological capabilities) are not the only important factor. The needs, expectations and preferences of intended users are also important, as users’ acceptance of the behavior change app influences their app usage and engagement, and thereby indirectly affects its chances of effectiveness. The focus group interviews discussed in Chapter 4 investigated these aspects, and revealed that people in the target population see the value in a physical activity app. They indicated that they prefer apps that motivate or coach them, provide tailored feedback toward personally set goals, and that enable competition with friends through a ranking or earning rewards. These features were especially important to participants who did not meet the guidelines for healthy physical activity levels; sufficiently active participants expressed interest in detailed activity information, for example to see how they could intensify their training sessions or improve their athletic performances. Also, they were not willing to share accomplishments regarding physical activities through general social media, but rather in a private online community. These are important insights for the development of new physical activity interventions.

For both app reviews, an evident limitation is that they create a snapshot of the state of the art of physical activity apps at a certain time point. This means that the results that were valid at the time of the analysis could get outdated after a while, and regular replications of the study should ensure whether the results are still relevant. In addition, since the total offer of health and fitness apps has grown to unmanageable numbers, it is impossible to cover all existing apps. Instead, a search and inclusion strategy is needed to navigate through the available apps and focus on a subset. Depending on the search strategy and inclusion criteria, the set of apps may be biased. The app review described in Chapter 2 applied a relatively strict selection on the apps (e.g., requiring some form of individual tailoring to the user), which could explain why slightly higher numbers of behavior change techniques were identified in the app analysis compared to similar studies (Breton et al., 2011; Cowan et al., 2013). Another limitation of both reviews is that the results depend on the recognizability of the behavior change techniques (in Chapter 2) or the technological features (in Chapter 3), as well as on the reviewers’ interpretation. Therefore, it is important to carefully design and operationalize the scoring framework, thereby ensuring high inter-rater reliability and replicability of the results.

One of the strengths of the focus group interviews is that the results are based on opinions of people from the target population of the Active2Gether system (i.e., Dutch young adults). Therefore, their opinions can be incorporated to the design fairly directly, without having to take possible differences in user preferences between subpopulations into account. However, this also implies that the results cannot be generalized to beyond the interviewed subpopulation. In order to establish whether the findings hold for other groups as well, the study should be repeated with participants from more diverse backgrounds. In addition, enriching the qualitative findings with empirical data of app usage should reveal whether these user preferences actually transfer to app engagement and adherence.

Altogether, the three studies presented in Part II contribute to a deeper understanding of the requirements for a new generation of app-based physical activity interventions. The results show that it is important to make better use of available knowledge and technology: on the one hand, incorporating more behavior change techniques (that are associated with
effectiveness), and on the other hand, implementing more technology (that enables smarter and more tailored support). Newly developed interventions that comply with these findings arguably will progress toward the look and feel of a virtual personal coach. That is also exactly one of the preferences expressed by intended users, and therefore also likely increases the probability of users’ engagement in the intervention.

17.2.2 Research question 2

What role can dynamic computational models play in the development of an intelligent mobile intervention for physical activity?

Computational modeling is a technique from artificial intelligence and computational science that is often used to study, predict and better understand behavior of complex systems. It requires knowledge of the domain under investigation, and an iterative approach of design and analysis of the modeled relationships (see also Section 1.4.3). In this thesis, several steps were undertaken to investigate the potential role of a computational model in a mobile intervention for physical activity.

First of all, Chapter 5 proposes a computational model of the process that the behavior change intervention is aimed to influence, i.e. the interplay of cognitive and social concepts affecting physical activity behavior. For the creation of this model, it was necessary to gain a thorough understanding of the literature on this process (Bandura, 1998, 2004), in order to be able to translate the knowledge into a formal specification. The resulting model was used to run simulations of many different scenarios that cannot be manipulated easily in reality. For example, the simulations show that for an active person, the perception of higher impediments leads to a decrease in physical activity. This effect is more apparent for a high value of the parameter that describes the effect of the impediments on the behavior. On the contrary, the presence of impediments causes a boost in the behavior of an inactive person. This effect is greater for a low influence of the impediments on the behavior. This suggests that encountering (and overcoming) obstacles could in some cases help to build confidence and as a result lead to an increase in physical activity. This is a good example of how carefully conducted simulations can lead to insights in the modeled behavior.

In addition, the computational model was analyzed by means of specification and analysis of properties that express dynamic patterns that are expected to emerge (Bosse, Jonker, et al., 2009). This analysis was done to automatically check whether the model behaves correctly and to test certain hypotheses, by running a large number of simulations and verifying such properties against the simulation traces. This process would be very time consuming if done by hand, and the use of automated property checking also enables quick verification of complex properties that would require deep logical thinking. For example, automated property checking revealed that all simulations of inactive persons who experience facilitating circumstances (and no impediments) show an increase in physical activity over the course of the simulation trace.

Although the model analysis or verification and expert validation discussed in Chapter 5 are valuable first steps in the evaluation of the presented computational model, the value of the model was further strengthened by validating the output of the model based on empirical data. This was done in Chapter 6 by testing the accuracy of the model to predict changes in physical activity levels over a period of two to twelve weeks. The predictions of the model were compared to empirical physical activity data of 108 different participants, in this case
the daily numbers of steps measured with a Fitbit One activity monitor. The results of this validation study show that the computational model performs rather well in predicting the changes in physical activity levels. The predictions showed a weak to moderate positive correlation with the actual data, which was statistically significant \((p < .05)\) for all predicted weeks. In contrast, a simple alternative model, based on randomly predicted end values for the physical activity behavior, had both weaker and non-significant correlations with the empirical data. Still, the random model performed better on some of the weeks than on others. This indicates that some characteristics of the data could have made it easier to predict the change correctly, and further investigation should find out why this happened. Nonetheless, the validation study presented in Chapter 6 contributes to the trustworthiness of the computational model that was part of the reasoning engine of the Active2Gether system, and therefore is a valuable step towards a more reliable and effective behavior change intervention.

However, when applying a computational model in a behavior change system, not only the goodness of fit matters. In the Active2Gether system, the model was incorporated in its reasoning engine, as a mechanism that predicts the most promising coaching strategy. (See Chapter 14 and maybe Section 17.2.4 for more details on the practical application of the model.) While the validity of the model’s predictions is undeniably essential, diversity in the content of the selected coaching messages is also important, in order to ensure the users’ engagement. Therefore, Chapter 7 describes an endeavor to find a set of parameters through a parameter tuning algorithm (simulated annealing), that yields values that stay close to the values based on literature but increases the diversity of the model outcomes as well. However, more importantly than finding this solution to the practical problem that arose when turning conceptual ideas into concrete implementations, the case study showcases a new application of parameter tuning in the field of computational modeling. The results from various simulation experiments with the model led to new insights in its behavior, and the global patterns of the resulting parameter sets provided information about (or could be explained by) the structure of the model and the meaning of its concepts and relations. This novel application of parameter tuning techniques is a new approach in the toolset of computational modelers, and therefore constitutes a valuable scientific contribution as well.

One of the limitations of the computational model presented in Chapter 5, and ultimately applied in the Active2Gether system as described in Chapter 14, is the fact that its design relied on a number of assumptions. Although the concepts, relations and parameter values of the model are based –where possible– on available literature on this topic (Bandura, 1998, 2004; Dzewaltowski et al., 1990; Petosa et al., 2003; Rovniak et al., 2002), formalizing the theory inevitably implies making decisions and simplifications to fill in the gaps (e.g., the numerical representation of the psychological concepts, the relationships as differential equations). This is not uncommon in computational modeling of complex behavioral processes, but it does mean that verification and validation of the model is important. The automated property verification and the expert validation in Chapter 5 provide initial evidence for the correctness of the model and face validity, and the study described in Chapter 6 takes the validation to the next level by comparing the model’s predictions to empirical data. Still, the presented model validation also has its limitations. The analyses reveal that the model performs rather well in predicting changes on a weekly level, but they do not disclose whether it also correctly represents the behavior on a more detailed (i.e., daily) level. Also, it would be relevant to find out whether the results are transferable to longer
periods than the currently tested twelve weeks. The model’s performance on predicting the underlying psychological constructs would be an interesting further exploration as well. In terms of validating this model for its application in a behavior change system, an evident improvement of the current study would be to investigate whether the model also correctly predicts changes in physical activity, while taking the effect of a certain coaching strategy into account. If simulation results including such effects are proven valid, the reliability of the decisions of the reasoning engine also increases.

The main limitations of the parameter tuning study presented in Chapter 7 are concerned with the fact that the results are based on only one case study, with one model, one evaluation measure (and related cost function) and one parameter tuning algorithm. Therefore, further investigation should reveal whether this novel use of parameter tuning techniques to study model behavior is also successful when extended to other applications.

Overall, in order to provide an answer to research question 2, Part III provides an in-depth exploration of a computational model for psychosocial influences on physical activity behavior, with a special focus on how to apply it in a real-life behavior change system. Initial steps towards validation of the model show promising results, thereby justifying its incorporation in the reasoning engine of the Active2Gether system. Applying a parameter tuning algorithm to increase the diversity of the simulation outcomes further consolidates the model’s suitability for application in the reasoning engine.

17.2.3 Research question 3

How can an individual’s social network be used to influence his/her physical activity behavior?

As explained in Chapter 1, social processes play an important role in achieving and maintaining a healthy lifestyle (Zimmerman and Connor, 1989). These processes rely on several motivational mechanisms, such as social norms, observational learning, social facilitation, social support and social comparison (Bandura, 1998; Buunk et al., 2013; Cheng et al., 2014; Festinger, 1954; McNeill et al., 2006). Some of these concepts were incorporated in the computational model that was discussed in Part III. The chapters in Part IV describe the role of several other social aspects in behavior change for physical activity.

First of all, Chapter 8 investigates the effect of changes in the structure of a social network on the diffusion of emotions through this network. Although the spread of emotions might not seem directly relevant for physical activity promotion, the mechanisms studied in this work also apply to attitudes and behaviors (Christakis and Fowler, 2013), and are therefore interesting foundations for network interventions that aim to stimulate behavior change. In that context, Chapter 8 proposes a method for finding effective network interventions to influence specific individuals. That method analyzes the structure of the network around the targeted individual to find strong transitive connections to people with a negative influence and weak transitive connections to people with a positive influence, and thereby determines which ties should be strengthened or weakened to achieve the most optimal effect on the targeted individual. The effect of these interventions was analyzed by simulating the diffusion of emotional values (e.g., about intentions and goals) through the social network, based on a model of social contagion (Bosse, Duell, et al., 2009). The simulation experiments demonstrate that it is possible to design a behavior change mechanism that influences specific
persons by affecting the social interactions between people and themselves. The simulations show that changing connections closer to the targeted individual yield a larger influence than changing connections further from the target, even if such further connections are the strongest or weakest links in a chain. A comparison of the effect of the proposed interventions with all possible interventions shows that they are among the most optimal possible interventions, i.e. in the 98th or 99th percentile. In addition, it was shown that targets with fewer connections are easier to influence than highly connected individuals. The proposed method provides support for the feasibility of behavior change systems that aim to support specific individuals by altering (online) interactions between people in their social network.

The proof of concept presented in Chapter 8 was based on simulations of a computational model of social contagion (Bosse, Duell, et al., 2009), applied to partially hypothetical data of a social network. After all, the structure of the network was adopted from the classical Zachary karate club network (Zachary, 1977), but the weights of the connections as well as the initial ‘emotion’ values were assigned randomly. Therefore, it is important to assess whether one of the most important building blocks of these experiments (i.e., the computational model) is valid, by applying it to real data and evaluating its accuracy. The work in Chapter 9 provides a first step in that direction. Here, physical activity data was collected for a network of 20 people over a period of 30 days. Based on the first seven days of data, the model predicted whether each participant’s physical activity level would increase or decrease, and the predictions were compared to the actual data. The results show that the contagion model predicted the direction of the change correctly in 80% of the cases. If only participants with significant trendlines in their physical activity data were considered, this performance went up to 87%. In addition, the mean squared error of the model’s predictions was close to the error of trendlines of the data (0.4272 and 0.3613, respectively), and much lower than the error based on extrapolations of the first seven days (6.8992). Considering that the model’s predictions are only based on seven days of data and the trendlines are derived from the complete dataset, we conclude that the model is able to predict changes in physical activity levels in a social network rather well.

Another relevant question regarding social processes in physical activity promotion is whether participation in social components of physical activity interventions influences the effectiveness of such interventions. Previous research had shown that partaking in the online community is associated with a higher physical activity level (Groenewegen et al., 2012), but the question remained whether it also makes a difference in the effect of a promotion program. In Chapter 10, this was investigated by analyzing a dataset of 50,000 people participating in a corporate health program. Approximately 5,000 of these participants chose to join a built-in online community to share their achievements. From that dataset, ten connected components were selected, together with a number of non-connected users (who did not join the community) with otherwise similar characteristics. The results demonstrate that the physical activity level of people that were willing to join a community showed an increase that was significantly greater compared to the other users. Since the data sets were balanced for possibly confounding factors like gender, time of the year and corporation, it is very likely that people’s willingness to become member of the community was the dominant factor that made a difference for their increase in physical activity. Although this study does not prove that an online community increases the effectiveness of a physical activity promotion program overall, it does show that it makes a difference for users who want to
participate in such an online social network. Further research should reveal whether actual participation in a community (rather than willingness to participate) has a similar (or smaller or greater) impact on the program’s effectiveness. Also, by assigning users randomly to each of these two conditions, the possibility that the differences are caused by other underlying personality characteristics should be excluded.

Although the analysis in Chapter 10 shows that the willingness to partake in an online community is associated with an increased effectiveness of a physical activity promotion program, it does not explain what the cause of this effect is. Therefore, the work in Chapter 11 builds upon the earlier findings, by investigating whether the computational model of social contagion (previously applied in Chapter 8 and studied in Chapter 9) is able to capture the dynamics in physical activity levels for users who were part of the online community of the corporate health program. In order to do so, we compared model predictions of such a contagion model (enriched with an expected linear increase as result of participating in the intervention) with predictions of a simple linear model to a dataset of 2,472 interconnected users. The results show that the enriched social contagion model performs better at describing the pattern seen in the empirical data than the linear model, both in terms of a significant difference in mean absolute error and correlation with the empirical data. These results indicate that some of the dynamics of the physical activity levels in the network could be explained by social contagion processes. This is vital information for designers of health interventions with a social component, as such models can then be used to maximize the benefits of social influence.

The preceding endeavors were all more or less concerned with social contagion processes. In addition to that, this thesis also explored social comparison as mechanism to influence behavior change for physical activity. Chapter 12 aimed to investigate the effect of showing users their preferred direction of social comparison (i.e., upward or downward) or showing the opposite direction. The results demonstrate that showing social comparison in the direction that users do not prefer is counter-effective. Therefore, this study shows that it is important to know users’ preferences: if not to enhance motivational effects, then at least to avoid adverse effects of showing social comparison that discourages them. Additional striking outcomes are that significant effects were found based on very small sample sizes, and that participants seemed able to indicate their preference reasonably well by simply answering one dichotomous question. Especially in light of the observation that many existing physical activity promotion programs implement social comparison in an upward direction, the results of this study argue that designers of such interventions should be aware of the risk of presenting a one-type-fits-all operationalization of social comparison to their users.

This work on answering research question 3 has some limitations. First, some analyses were done based on hypothetical data. This is the case with the simulation experiments of social network interventions in Chapter 8 (where the connection strengths and emotion values were assigned at random), and with the simulations of the contagion model in Chapter 11 (where the connection strengths were based on generic values). The parameters of the social contagion model in Chapter 9 were derived from real data, but were based on heuristics. Validation of these heuristics would further strengthen the reliability of the outcomes of that study. Second, in both studies in which the accuracy of the social contagion model was tested (described in Chapter 9 and Chapter 11), certain assumptions implied limitations on the findings. In Chapter 9, social contagion was the only factor considered as
Chapter 17. Discussion & Conclusion

Influence on the physical activity behavior in the group, and Chapter 11 only considered connections in the online community (and not in real life) in the social contagion process. Both simplifications imply that the models described an incomplete picture of the social processes taking place. A third limitation has to do with the generalizability of the results. For example, the analyses in Chapter 10 were based on whether or not users joined an online community at some point in time, but not necessarily during the same period as the analyzed data. Therefore, it is important to be very careful when interpreting the results and when drawing conclusions about the effect of such social features in general. In addition, the results of Chapter 11 cannot easily be transferred to the general population, as the analyses are based on data collection in the context of a physical activity promotion program. Similarly, the findings in Chapter 12 were based on a small sample of male amateur soccer players, so the results should be interpreted with caution. Finally, it is important to realize that the research presented in this thesis (although it provides valuable insights on social aspects in behavior change interventions) only scratches the surface of the vast amount of social processes in behavior change. Further research should continue studying these and other mechanisms in more detail.

In conclusion, the work presented in Part IV provides more insight in the importance of social processes in physical activity interventions, thereby contributing to answering research question 3. It was shown that users who choose to join an online community (at some point in time) benefit more from a physical activity promotion program than users who do not participate in the community. Also, a computational model of social contagion was supported by the results of two different data analysis studies. In addition, this part contains an investigation of how to exploit such processes in behavior change systems, specifically by altering connections in a user's social network to steer social contagion or by selecting other users for social comparison while taking the user's preference into account. This provides important guidelines for developers of behavior change interventions.

17.2.4 Research question 4

To what extent can the answers to the questions above be used to design, implement, exploit and evaluate a personalized mobile intervention for physical activity promotion?

The previous research questions each highlight parts of the multi-faceted subject of innovations in behavior change for physical activity. This research question, addressed in the chapters in Part V, deals with bringing all components together for the design, implementation, exploitation and evaluation of a personalized mobile intervention for physical activity promotion.

First of all, Chapter 13 describes the step-by-step approach to designing the Active2Gether intervention. This consists of defining the intervention's primary and secondary objectives, establishing the theoretical framework, developing the content of the intervention, pilot testing the intervention, employing the intervention, and finally evaluating its effectiveness. Planning and reporting this development process meticulously stimulated approaching it in a structured manner, thereby ensuring that the applied techniques were grounded on evidence and well-integrated into one overall reasoning engine. The resulting chapter serves as a guideline for the design of behavior change interventions, by providing a starting point. Also, the detailed documentation increases the transparency of the current intervention, which helps to interpret the results of studies on its effectiveness and user
17.2 Discussion of research questions and results

Whereas Chapter 13 summarizes the components and reasoning mechanisms of the Active2Gether system, it does not provide details on its implementation. Therefore, Chapter 14 describes the technical development of the Active2Gether intervention and its underlying system, together with a reflection on the technical design choices. A recurring theme in the lessons learnt is the need for more flexibility than the system currently offers. To name a few examples, more flexibility would produce an improvement for the data collection (i.e., relying more on unintrusive measurements and less on user input), for the implementation of the ranking feature (i.e., less strict application of the user’s preference for upward or downward social comparison) and for the suggestion of a coaching domain (i.e., considering all activities related to a certain physical activity domain, rather than only activities associated with the reported significant locations). Also, the chapter suggests some ideas to make the coaching messages more personally and contextually relevant to the users, which we believe would improve the system’s effectiveness and user experience as well. As with the intervention design described in Chapter 13, such comprehensive documentation may serve as a source of inspiration or point of reference for the development of other behavior change systems. Additionally, the lessons learnt help others to identify potential pitfalls and opportunities to improve upon our system.

The previous two chapters provide insights into the design and implementation of the Active2Gether intervention. The intervention was employed in a user study to assess its effectiveness and to investigate the user experience. In Chapter 15, a study of the effectiveness of the Active2Gether system in increasing the physical activity levels of the participants is presented. The results show no significant effect on the physical activity levels in the two intervention groups (Active2Gether-Full and Active2Gether-Light), and no significant change in the underlying behavioral determinants either. Still, a clinically relevant increase of 4.2 to 4.4 daily minutes of moderate to vigorous physical activity was observed, which does suggest that there was an upward trend in the physical activity behavior. Therefore, it would be interesting to repeat the study over a longer period of time and with larger sample sizes, in order to see whether a significant effect would be found in more optimal circumstances.

For app-based interventions to be successful in supporting users to achieve or maintain healthy behavior, not only the effectiveness of the app is important. The users’ appreciation of the app should not be overlooked, as it can influence engagement with and adherence to the app. In Chapter 16, we explore the user appreciation of the Active2Gether intervention. The results show that the participants wanted a coaching feature to be included (on top of self-monitoring functionalities), but were also critical of how it was implemented (in terms of number and content of the messages). The coaching should be perceived as personal and relevant, and it should be sufficiently diverse in order not to become too repetitive. Thus, a (personal) balance in the number of messages should be found: too many messages seemed to be annoying, but on the other hand, such system-initiated user interaction appeared to reduce dropout. Overall, the user evaluation of the Active2Gether system (in comparison with the commercially available Fitbit app) helps to understand which aspects were appreciated by end users. This information could help to increase engagement, and thereby adherence and overall effectiveness. The fact that this evaluation was based on real use of the apps for a considerable amount of time, rather than based on hypothetical use scenarios, makes it an even more valuable contribution to research on behavior change applications.
During the process of designing, implementing and evaluating the Active2Gether system, several opportunities for improvement came to light. Unfortunately, the user study was not able to prove that Active2Gether was effective in increasing the users’ number of active minutes. This does not necessarily mean that the underlying design principles are faulty, but maybe that they were not executed well enough to see significant effects. The reflections on the technical design choices in Chapter 14 provide several suggestions to improve the current system. In addition, the user evaluation study reported in Chapter 16 also indicates how the user experience could be improved. For example, the system should be more technically stable, the physical activity data in the dashboard should be more detailed and updated more often, and the coaching should be more personally and contextually relevant. It is promising that, despite the small sample sizes and the short exposure to the intervention, small but clinically relevant changes in daily number of minutes spent on moderate to vigorous physical activity were observed. Repeating the user study, after implementing the identified improvements in the Active2Gether system, with larger sample sizes and over a longer period of time, should show whether it is indeed a worthwhile endeavor to incorporate evidence-based techniques and methods from artificial intelligence to improve mobile behavior change interventions.

Altogether, the chapters in Part V provide a complete description of the process of developing and evaluating an innovative mobile behavior change intervention for physical activity. Where relevant and possible, the results from the research presented in the preceding parts were incorporated into the system’s design. Also, several improvements for the development of such systems were suggested. Therefore, this part contributes on the one hand by proposing a design for a behavior change system with innovative elements, and on the other hand by providing hands-on guidelines for future work in this domain.

17.3 Ethical aspects

When using technology to monitor and influence people’s behavior, it is important to take ethical considerations into account. After all, a considerable amount of personal data is collected and used to derive tailored feedback that is aimed to influence the user’s behavior. In order to respect the users’ privacy and autonomy, this should be done carefully and with high regard for ethical implications. Therefore, this section reflects on some ethical aspects that are relevant with respect to the design and employment of the Active2Gether system.

The reflection is based on the eFRIEND ethical framework for intelligent environments development (Jones et al., 2015). This framework is a combination several other frameworks addressing ethical issues, and extends it with previously overlooked items. It is focused on so-called intelligent environments, i.e. context-sensitive services to humans in the physical space. Although the Active2Gether system does not fall exactly within the definition of an intelligent environment (Augusto et al., 2013), the ethical principles are largely similar. Each of the following subsections addresses one of the general principles from the framework (Jones et al., 2015), and a brief summary is presented at the end of the section.

17.3.1 Non-maleficence and beneficence

The first ethical principle concerns the non-maleficient and beneficient purpose of the system. Clearly, the Active2Gether system was designed to increase the users’ welfare and quality of life by encouraging healthy behavior. By supporting users to choose a beneficial coaching
domain and to set a realistic goal, the system proactively offers opportunities to assist the users. The current implementation of the system does not include a mechanism to detect extreme physical activity levels and to advise users to slow down, but the system’s architecture does allow implementing this in a next version. Still, the Active2Gether system is built in such a way that it stimulates physical activity in the users’ daily life, and it does not reward extreme physical activities. Also, new goals are suggested in small increments, and the reasoning engine allows the users to stop being coached towards higher physical activity levels once they have reached the norms.

17.3.2 User-centricity

Another ethical principle is that the users’ wishes should be central in the development process. In order to develop a behavior change system that meets the users’ needs and wishes, focus group interviews were conducted to inform design choices (see Chapter 4). Also, as described in Chapter 15, the system was first pilot tested within the Active2Gether team and with people from the target population, in order to solve any technical issues and to tweak some functionalities. This way, we attempted to keep the user in the design loop. Still, the development process would ideally have had more iterations with the users. The findings from the evaluation study (in Chapter 16) provided helpful starting points for improving the system, and one can only imagine that more feedback moments would lead to further improvements of the intervention.

17.3.3 Multiple user groups

The principle of multiple user groups is related to the situation in which an intelligent system is used by multiple people at the same time. In case of intelligent environments, this is a valid consideration. However, as the Active2Gether system is designed to be run on the smartphone for one individual, this ethical aspect is not relevant.

17.3.4 Privacy

A term that is unquestionably brought up when discussing ethics of technology, is privacy. In a system like the Active2Gether app, privacy could be ensured by letting users decide the level of detail in which data is collected. This was taken into account during the development of the social comparison feature. For people granted permission to view their Facebook connections and who are connected to another Active2Gether user on Facebook, their full name is shown in the overview. Users who don’t know each other or who did not allow the Active2Gether system to look up their Facebook connections, only the first two letters of their first name are shown, thereby maintaining a level of anonymity.

An evident example that would benefit from more flexible is the location monitoring, which could be turned off or set to other time intervals or accuracy levels. Unfortunately, the current implementation of the Active2Gether system does not support such custom settings. Although users can turn off the location monitoring, this will undeniably deteriorate the quality of the coaching as well as the user experience. Similarly, physical activity data could also be stored at higher levels of aggregation, according to the users’ preferences. The privacy aspect is therefore an important focal point for improving the ethical aspects of the Active2Gether intervention.
17.3.5 **Data protection**

A related ethical concern is data protection. It involves the users being in charge of their data and information sharing, as well as informed consent to the use of personal data and good practices of data protection. The Active2Gether system did show the physical activity data on the dashboard, and a separate personal profile page showed some additional background information, such as the current coaching domain, the current goal and an overview of the reported significant locations. The users could, for example, change their current goal, but changing information about their significant locations required involvement from the team. Also, deleting their personal data or their complete account was not possible for the users themselves. Such functionalities should therefore be implemented and/or professionalized before employing the system in real-life scenarios.

Still, Active2Gether adhered to recognized principles of data protection. This was documented in a study protocol before commencement of the final study, and it was approved by the medical ethical committee of the VU medical center. Also, all participants were informed about the data collection, and they signed an informed consent before being admitted to the final study. When employing the Active2Gether in real-life settings outside academic purposes, such a procedure should still be in place to fully inform users of the personal data that is stored during their use of the system.

17.3.6 **Security**

Another ethical aspect related to privacy and data protection is the system’s security. It concerns the need to protect users and their information, the reliability and stability of the system and the security of the data transfer. In the Active2Gether system, unauthorized access to the users’ personal data is avoided by protecting all accounts and dashboards with a password. In addition, safe data transfer is ensured by making use of a secure HTTPS connection. Also, personally identifiable information (including the users’ email addresses and first names) and passwords are encrypted, using either a MD5 hash function or AES encryption (Daemen and Rijmen, 2013; Rivest, 1992). However, when employing an intervention like Active2Gether in real life, data that could be used indirectly to identify users should also be encrypted, in order to reduce the risk of any security breaches.

Although there were no signs of the Active2Gether system being unreliable or unstable during its use in the final study, more mechanisms should be in place to ensure its stability at potential higher demands.

17.3.7 **Autonomy**

The Active2Gether system aims to support its users’ in their efforts to achieve behavior change, while respecting their independence and autonomy. An evident example of this intent is the fact that the reasoning engine presents the users with suggestions, which can be overruled by them. For example, based on the physical activity data and context information, the system suggests a coaching domain to focus on, but the users decide whether they want to follow the advice or not. The same holds for setting a goal: the system does suggest a goal based on the user’s physical activity data, but the user can easily adjust it before confirmation.

Still, the user does not have unlimited influence. In the current implementation, it is not possible to put the system ‘on hold’ (i.e., not receive any coaching messages and/or
questions). It is imaginable that such a feature would be relevant if the user’s regular daily life is interrupted, for example when he/she is ill or on holidays. In addition, as suggested in some of the previous sections, the user’s autonomy would be further enhanced by enabling more custom settings for data collection and information sharing.

17.3.8 **Transparency**

Unlike with some intelligent environments, users of the Active2Gether system do not have to be given notice of the existence of the system, since it is impossible to be unaware of it. However, it would be recommended to inform users about the system’s capabilities, in order to shape their expectations and to inform them of any potential weaknesses. For example, as long as the system does not include a mechanism to detect extreme physical activity behavior, users should be pointed at their own responsibility to behave responsibly. This also holds for the other end of the scale: the system aims to support users to live a healthier life, but people in need of medical intervention should not rely on the Active2Gether system alone, as that goes beyond the aim and capability of the intervention.

In terms of data collection and processing, the Active2Gether system is fairly transparent. Since a substantial part of the intervention relies on self-monitoring, much of the collected data is relayed back to the user via the dashboard. That way, users are also made aware that this data is being collected. This also holds for the location data, although to a lesser extent. This data is not shown to the user, but it is used as trigger to ask the users questions about their travel modes to certain locations. Because of that, the users should be aware that their location is being monitored in the background. Also, all users have to grant permission to use their smartphone’s location services, so it never happens unnoticed. In case the user input on travel options would become obsolete because of other reliable methods for collecting this data, it would be recommended to implement a feature that reminds the user of this data being collected (e.g., showing the location data on the dashboard as well), in order to increase the system’s transparency.

One aspect of the Active2Gether system that could benefit from more transparency is the dashboard panel that shows a ranking of the weekly number of steps of the user and six other users. The selection of these other users is based on an algorithm that places high priority on the user’s preference for the direction of social comparison (cf. Chapter 12 and Chapter 14). This implies that users are often found at the top or bottom position of the ranking, regardless of (changes in) their physical activity behavior. The basis for this selection mechanism was not transparent to the users, and was also sometimes perceived as demotivational, as efforts to increase physical activity were not rewarded by a higher position in the ranking.

17.3.9 **Equality, dignity and inclusiveness**

The final ethical principle considers equality, dignity and inclusiveness. There is an ethical consideration with respect to the intended user groups. As argued throughout this thesis, the Active2Gether system focuses on young adults (aged between 18 and 30 years) and is only available for smartphones running on Android. Therefore, many of the intermediate studies and the final study were conducted with participants from this age group. In practice, many participants were highly educated and/or university students and living in urban areas. Consequently, the design of the Active2Gether system is optimized for this population, and might be less effective or appealing for other groups. This is an undesirable consequence,
especially since people with a low socioeconomic status are overall less physically active than the general population (Giles-Corti and Donovan, 2002), and they are harder to reach with health interventions (Iliffe et al., 2017). Ideally, the Active2Gether system should be extended and tested for other population groups and operating systems as well, so it can be equally beneficial for every potential user.

On the other hand, the system is inclusive in the sense that it is (potentially) very affordable. In principle, the intervention does not require any human involvement, and is therefore quite cheap to maintain and employ for many users. Therefore, the cost for downloading and using the Active2Gether app could be kept at a minimum (or even free). Currently, the Active2Gether system does rely on measurements from a Fitbit device, which might be a financial burden for some potential users, but an extension of the system to allow using measurements from the smartphone itself is feasible.

Summary
Overall, the Active2Gether intervention respects the ethical principles from the eFRIEND framework well. However, there is still some room for improvement if the intervention would become available to the public outside the academic setting. For example, more data stored on the system’s server should be secured by means of encryption, more custom settings for data collection and information sharing should be available to the users, and the intervention should be made suitable for a broader audience than the current target population.

17.4 Future work
The research presented in this thesis explored several aspects of using technology for behavior change support. Especially in such a relatively new discipline, there are always many opportunities for further investigation. This section addresses a number of promising directions for future research to follow up on this thesis.

First of all, the research into the requirements for mobile behavior change interventions (presented in Part II) is inevitably linked to the current state of the techniques and science. As modern technology and understanding of human behavior formation progresses, new theoretical insights and technological developments lead to new user preferences, and consequently new requirements for such interventions. Therefore, such analyses should be regularly replicated, in order to validate or update the latest findings. In addition, there is a need for in-depth reviews of effectiveness of mobile behavior change interventions. This would contribute to understanding which mobile behavior change interventions (or on a more detailed level, which behavior change techniques) are successful in effectuating behavior change. If sufficient data is available, such analyses could also reveal which type of intervention or technique is effective for which type of user. That way, a new generation of interventions could be designed in an even more tailored fashion.

With respect to the role of computational models in the development of mobile interventions for physical activity, additional validation studies of such models are recommended to further consolidate their reliability. Naturally, more reliable and realistic computational models imply to more veracious new insights in human behavior through simulations, as well as more effective behavior change support systems. Also, it would be interesting to compare the validity of computational models of behavior change based on different
theoretical frameworks. In addition, it is highly plausible that there are individual differences in the psychosocial and behavioral processes represented by these computational models, which can be characterized by differences in their parameter values. It should be investigated whether this is indeed the case, and if so, individual differences should be taken into account. These individual parameter values could either be estimated based on responses to (personality) questionnaires, or learned automatically from behavioral data. Allowing individually set parameters could then lead to better performance of the models and more truth value in their simulations.

Another interesting direction for future research in the domain of computational models of behavior change processes is related to the way they are generated. Even with parameters tuned individually based on user data, a computational model as presented in Part III follows a top-down approach. It is designed based on theory, which means that it makes of existing knowledge and that its internal dynamics and simulation outputs are interpretable. However, with the increasing availability of relevant data, it is interesting to investigate more bottom-up approaches as well. In other words, such models could be learned from data. This does require a large body of data from a large number of participants, but it could lead to unforeseen insights and potentially higher validity. Apart from building models of behavioral processes from data, similar methods could be used to directly derive prediction mechanisms for determining coaching actions of behavior change systems. If such approach is successful, this could render the use of computational models in reasoning engines of behavior change systems unnecessary, but at the same time, it could also lead to new insights that can be incorporated in top-down generated computational models.

As mentioned before, many social processes have been established to play a role in behavior formation and behavior change. The work in this thesis only scratches the surface of this area of research, so further research into these and other social influences should be conducted to investigate which processes could be beneficial (or detrimental) to people’s attempts at behavior change. On the one hand, such investigations would contribute to further building theory on this topic, by gaining new insights in human behavior. On the other hand, the research should also focus on development of behavior change interventions, by exploring how these processes could be applied to help change behavior and how the underlying principles could be operationalized.

Another valuable opportunity for future research concerns the further development of the Active2Gether system or similar innovative behavior change interventions. This thesis described the design of such an intervention in detail, and provided many ideas for improvement as well, for example in Chapter 14 and Chapter 16. In addition, the previous section discussed some guidelines for better adherence to ethical norms.

One of the main directions for potential improvement of the Active2Gether system is related to the augmentation of the ‘intelligence’ in its reasoning engine. Incorporating more intelligent methods could make the system more flexible and adaptive to the users, which would contribute to the intervention being perceived as a virtual personal coach. Such intelligence could be achieved by incorporating more data-based methods and using feedback mechanisms to update the coaching offered to the users. Feedback could be obtained directly from the users (e.g., by letting them indicate whether they liked a certain coaching message), or from user data about their behavior or underlying psychological determinants (i.e., by deriving whether some action had a visible effect). This way, the system could adapt to each individual users and to the users’ progress over time.
In addition, if sufficient data is available, principles used in recommender systems could be applied. For example, collaborative filtering is an algorithm based on the assumption that if two users have a similar opinion on something, then the opinion of one of those users on something else is a good predictor for the opinion of the other user. In other words, if a certain coaching message is effective for one user, and this user is similar to another user, then the message will probably also work for the other user. However, such approaches require a vast amount of data over a longer period of time, and probably a less complex system, in order to be able to straightforwardly link observations of effectiveness to specific messages or behavior change techniques.

Another approach to more intelligent coaching relies on more contextual awareness and personal relevance of the messages. For example, geo-fencing (i.e., monitoring a virtual perimeter in a geographic) could be used to trigger sending certain coaching messages, for example when a user arrives home. In addition, the coaching messages could incorporate more information about the users’ behavioral patterns, such as the days when they usually exercise or the time of the day when they usually leave home on workdays. This again could be learned from user data by automatically recognizing patterns in time series of their daily activities. Also, direct triggers to send certain coaching messages from the user data could be implemented, such as when reaching a daily or weekly goal. This does not necessarily require very sophisticated methods, but it could improve the perception of the intervention as a virtual coach. Also, such messages (that are triggered by locations, behavioral patterns or physical activity data) are likely to contain practical and personal feedback, and thereby increase engagement with the intervention. The emergence of the Internet of Things (IoT) leads to the availability of many new types of (more detailed) data that are relevant for monitoring and influencing behavior (Gubbi et al., 2013; Whitmore et al., 2015). This development will likely further advance the use of data-based techniques in behavior change systems.

A key strength of the Active2Gether system is that it is based on an architecture that can be adapted relatively easily. Each of its building blocks (such as the mechanism for data collection, the reasoning engine, or the dashboard interface), can be improved, extended or completely replaced with alternative solutions. That way, it allows for flexible implementation of the improvements described above.

17.5 Conclusion

The research presented in this thesis investigates several aspects of using technology to stimulate behavior change for physical activity, and it proposes the design of an intelligent physical activity promotion app (i.e., the Active2Gether system). In doing so, techniques from (mobile) technology and artificial intelligence are applied, as well as scientific knowledge from human-directed disciplines within psychology and social sciences. The role of social processes in establishing (and maintaining) healthy behavior is considered in particular.

In Part II, the state of the art of mobile behavior change interventions for physical activity and user preferences of the target population are discussed, in order to gain insight into the requirements for such interventions. The results show that it is important to make better use of available knowledge and technology: on the one hand, incorporating more behavior change techniques (that are associated with effectiveness), and on the other hand,
implementing more technology (that enables smarter and more tailored support). Also, intended users expressed their preference for an intervention that has the impression of a virtual personal coach.

Part III discusses the role of computational models in the development of mobile interventions for physical activity. Together, the chapters provide an in-depth exploration of a computational model for psychosocial influences on physical activity behavior, with a focus on how to apply it in a real-life behavior change system. Initial validation studies of the model show promising results, thereby justifying its incorporation in Active2Gether’s reasoning engine. Applying a parameter tuning algorithm to increase the diversity of the simulation outcomes further consolidates the model’s suitability for application in the reasoning engine.

In Part IV, the role of social processes in establishing behavior change is studied. It is shown that users who choose to join an online community (at some point in time) benefit more from a physical activity promotion program than users who do not participate in the community. Also, the validity of a computational model of social contagion is supported by the results of two different data analysis studies. In addition, this part contains an exploration of ways to exploit such processes in behavior change systems, i.e. by altering connections in a user’s social network to steer social contagion and by selecting other users for social comparison while taking the user’s preference into account. These findings provide important guidelines for developers of behavior change interventions.

Part V provides a complete description of the process of developing and evaluating an innovative mobile behavior change intervention for physical activity. Where relevant and possible, the results from the research presented in the preceding parts were incorporated into the system’s design. Also, several improvements for the development of such systems are suggested.

Overall, the work presented in this thesis contributes to the scientific advancement of the domain of intelligent behavior change interventions for physical activity, by investigating several approaches to incorporate the use of technology in analyzing, understanding and supporting human behavior. In addition, this thesis presents practical steps and insights with respect to the development of such an intelligent physical activity promotion intervention. We hope and expect that this work contributes to the further development of sophisticated physical activity interventions, and thereby to a healthier society.
References


Breton, Emily R., Bernard F. Fuemmeler, and Lorien C. Abroms (2011). “Weight loss — there is an app for that! But does it adhere to evidence-informed practices?” In: Translational behavioral medicine 1.4, pages 523–529.


Giles-Corti, Billie and Robert J. Donovan (2002). “Socioeconomic status differences in recreational physical activity levels and real and perceived access to a supportive physical environment”. In: Preventive medicine 35.6, pages 601–611.


