Using extensive data in design and evaluation of BCSS

Floor Sieverink [1], Jobke Wentzel [1,2], Randy Klaassen[3], Robby van Delden [3], Dirk Heylen[3], Harri Oinas-Kukkonen[4] and Lisette van Gemert-Pijnen [1]

[1] Centre for eHealth and Wellbeing Research, Department of Psychology, Health and Technology, University of Twente, the Netherlands
{f.sieverink, m.j.wentzel,j.vangemert-pijnen}@utwente.nl

[2] Section Health Care, Nursing, Saxion University of Applied Sciences, the Netherlands

[3] Human Media Interaction, University of Twente, the Netherlands
{r.klaassen, r.w.vandelden, d.k.j.heylen}@utwente.nl

[3] Oulu Advanced Research on Service and Information Systems Group, Faculty of Information Technology and Electrical Engineering, University of Oulu, Finland
{harri.oinas-kukkonen}@oulu.fi

Abstract. The International Workshop on Behavior Change Support Systems provides a place to discuss recent advances in BCSS research. The aim of this year's workshop was to discuss and share examples and experiences from research on BCSSs regarding both the evaluation and improvement of these systems, as the systems use more and more types of data over longer periods of time. Fitting the multidisciplinary character of the workshop, a wide variety of topics was presented in which the use of extensive data plays a role: occupational health, diet and healthy foods, persuasive dialogues related to food intake, entertainment for people with disabilities, and telemonitoring of heart failure patients. Group discussions focused on two key aspects related to the use of more extensive data: approaches to encourage transparency of the use of data to prevent `black-box thinking', and pointing out a strategy for data driven coaching over longer periods of time.

Keywords: behavior change support systems; BCSSs; extensive data; persuasive design; evaluation.

1 Introduction

Behavior change support systems (BCSSs), with its specific implementations of technology, have the potential to persuade people to change behaviors or even lifestyles. With its ever growing possibilities, it is impossible to imagine our everyday life without smart technology. Technology is not only embedded into our homes or workplaces, people are carrying more and more technology with them throughout the day. Technology that truly blends in with our lives and habits, and adapts to measured preferences,

behaviors, and responses, may more effectively support our goals and steer our behavior. The increased pervasiveness of technology provides the opportunity to use real world data tracked in daily life over longer periods of time in persuasive systems [1]. These developments shed a new light over the design and evaluation of BCSSs [2].

The increased use over time of persuasive systems introduces a number of challenges, as well as opportunities. For example, the data that is being collected via BCSSs provide great opportunities for data-driven coaching to enhance and retain health and wellbeing of users. Large amounts of valid, real-time, context-based, and technology-generated data (e.g., log data [3] and monitoring data) allow for developing and advancing behaviour change models and theories to support individuals in changing their behaviour and to enable the personalization of technologies [4, 5]. However, this requires reliance on the validity of the used data, and on the choice of the timing, modality, content, and even the right device for the persuasive system. This poses an important challenge for research and the development of persuasive systems in different contexts, such as health [6], energy, or the everyday workplace.

At the same time, this increase of valuable quantitative data requires reflecting on the evaluation of BCSSs. Traditional evaluation approaches do not always provide insight into the unique interaction process between the user and the technology. In fact, a description of how the technology supports the user in performing certain behavior is often missing in such evaluations [7]. Furthermore, to accommodate to the complexity of behavior change, BCSSs often consist of multiple components that people can use in different ways in terms of what elements they choose to use, as well as the frequency, time, and place of use [8, 9]. Therefore, the experienced content can differ across all users. Subsequently, through its use, technologies are often tailored and personalized to the individual. Therefore, it seems current evaluations provide limited insight into the process of how the use of the different components of the technology has contributed to healthier living, improved wellbeing, or a user's ability to conduct daily tasks [7, 9, 10]. Thus, the characteristics of more pervasive technology and the influence of the user and the context in which the technology is implemented and used, change the way evaluations should be conducted [3, 11]. Therefore, it is necessary to develop and evaluate innovative approaches that allow for a close look into the process by which users find and share information, and a more holistic view on how users gain benefits out of the persuasive technology [12, 13].

The sixth International Workshop on Behavior Change Support Systems provided a place to discuss recent advances in BCSS research. The aim of this workshop was to discuss and share examples and experiences from research on BCSSs regarding the use of more and more types of data for evaluation and improvement of these systems. The following section lists the broader set of topics of interest of which some were represented in the submissions.

2 Topics/Themes

Topics for submissions included, but were not limited to:

- Smart monitoring for persuasive coaching especially in (but not limited to) the area of health and well-being
- Developing just-in-time persuasive prompts and feedback to support behavior and to create adherence and engagement to different technologies, using data generated by smart sensors, self-tracking devices, wearables, etc.
- Engagement, integration, connectivity, personalization, and changes in Persuasive Technology
- Interactive visualizations (including virtual coaches and dialogues) for personalization and social support
- High tech, human touch/humanizing technology
- Connectivity designs for social support, e.g. for lifestyle change and improving well-being
- Design guidelines for the design, implementation and evaluation of BCSSs
- Persuasive strategies related to different outcomes (engagement; resilience; attitudes; compliance; behaviors) and levels (individual; community; society) of change

This year, we especially welcomed papers regarding the evaluation of BCSS:

- Methods for measuring the impact of BCSSs and smart persuasive environments on individuals, community, and society
- Methods for measuring the effect of persuasive strategies on task adherence (e.g., via fractional factorial designs)
- Methods (including mixed methods approaches) for measuring various aspects of BCSSs in the wild; considering context and including process and product measurements in a real-life setting
- Methods or approaches to evaluate the persuasiveness of different technologies for BCSSs (mobile, ubiquitous, ambient technologies, virtual environments, sensor-based, etc.)
- Advanced big data analytics for analyzing and interpreting usage data and self-tracking data from (multimodal) sensors
- Translating the outcomes into multimodal feedback cues, and their effects on adherence and outcomes
- Advanced analytics to predict adherence and to identify usage patterns and its effects on adherence
- Implementation strategies to deal with proprietary closed algorithm layers to gather reliably gather data of daily use, using commercial sensor devices

3 This year's contributions

The five papers presented during this year's edition of the workshop all relate to the topic of how to use extensive data for the design and evaluation of BCSSs. At the core of any BCSS should lie a clear vision of how persuasion can occur. In the first paper of these proceedings, Kekkonen et al. contribute to this vision with a systematic literature review to gain insight into existing interventions in occupational health [14]. This review presents an overview of recognized persuasive features that can be used in the

development of occupational health support systems for behavior change. An important conclusion of the review is that monitoring and analyzing employees' health can be beneficial for both the employer and the employee. Although, it might also raise questions regarding the legal and ethical aspects when the aim of the employer for monitoring the employees behavior might differ or go beyond the generally welcomed improvement of the working conditions.

In the second paper of this proceedings, Manninen & Tikka describe the development process of a mobile system that stimulates the vegetable and fruit consumption of its users [15]. This BCSS incorporates gamification elements in combination with an Implicit Association Test (IAT) to guide users' eating behavior. In this game, data was collected to calculate scores indicating the players' attitudes regarding healthier food. These scores were used to provide the player insight into his or her thinking, hopefully also providing a first important step to change his or her (unfavorable) behavior.

Using automatically collected data, such as log data or observational data for actively improving a system or tailoring its response to specific user actions, requires reliable data measurement. In the third paper of this year's BCSS proceedings, Van Delden & Reidsma describe issues with the reliability and validity of data that can be at play when relying on automatic data collection and analysis [16]. In their study, simplified motion energy analysis was used in the evaluation of an interactive ball that responds to gross body movements and vocal interactions. However, when automatic analysis is not completely accurate, the system might render results that are not representative, persuasive, or motivating. For example, if a system compliments a user for an increase in (measured) physical activity, while in fact the user was sedentarily enjoying incoming rays of sun, receiving positive reinforcement will not contribute to his or her motivation to continue using the system. Thus, analyzing correct data, using thorough sensitive analysis, and acting correctly upon the data are crucial, and might become at risk when incorporating more extensive sets of data.

Once a technology is in use, researchers and developers can use extensive usage and monitoring data to evaluate the system and to better tailor interventions to the user. In the fourth paper, De Franco et al. propose an evaluation matrix for measuring the persuasiveness of BCSS by analyzing interview data [17]. They introduce three evaluation dimensions from argumentation (i) the type of persuasion, (ii) the person being persuaded and (iii) the type of evaluation, and show how these apply to a real world example of behavior change.

An application of using monitoring data to evaluate BCSSs is found in the last paper of this proceedings. In this paper Cruz-Martinez et al. describe how daily monitoring of patient measurements (e.g., weight, blood pressure, and heart rate, via logfiles of the use of a platform (IMediSense)) in combination with qualitative interviews and usability tests can be used to guide heart failure rehabilitation to prevent complications and to coach patients health and wellbeing at home in a personal and motivating fashion. At the hearth of this persuasive data-driven coaching approach is the premises that keeping self-monitoring as a simple process is important to promote adherence to the system among the users. Consequently, it is envisioned that the data analysis will enable the timely identification of vital symptoms to prevent complications (resubmission to a hospital) [18].

4 Group Discussions at BCSS 2018

After the work in the previously described papers was presented, the group of participants was split in two groups in order to discuss two main challenges during the second half of the workshop. We introduced nine different challenges to choose from, five were based on the papers that were presented and there were four more general challenges related to extensive data in the design and evaluation of BCSSs. For each challenge we also provided one example of how we interpreted the challenge.

The group of participants was split into two groups and invited to pick one or more challenges for a group discussion. During the group discussion, participants were invited to elaborate on how they interpreted the challenge from their point of view and background. As a second step they were ask to add their experiences and examples from their own research. Based on this, participants were invited to discuss potential approaches to deal with the related aspects of the challenges. To conclude the workshop, both groups prepared a pitch to present their preliminary conclusions regarding their challenges and the outcomes of the discussions. The discussed challenges and the outcomes are described below.

Challenge 1: How can we keep the BCSS transparent for people (both developers and users), considering black-box thinking; how to let the system tell why it does something? During the discussion it became apparent that transparency primarily is the ability to explain why things happen in a system. In BCSSs, transparency can lack in several stages of the research process and in different elements of a system or. It often remains unclear what (behavioral) theory the system is built upon, of what elements a system exists and how these elements were operationalized (both functional and in the design). For example, how does a BCSS that uses data for automatic coaching make decisions? What parts are based on adaptive/learning algorithms, and what decisions are actually made by the designer? The participants agreed that this lack of transparency can have negative consequences for the trust in and use of systems. Several approaches were suggested to tackle this negative consequence. First, it was stated that true transparency goes beyond merely a description of the BCSS from an expert (or technological) point of view. It also requires the effort to support the `average user' in understanding the working mechanisms of a system. In that sense, the advice of the participants to researchers and developers would be to not *" celebrate the mystery of the algorithms"*. Rather, it is essential to explain the working mechanisms of a BCSS on different levels, in such a way that it is also understandable for the general public.
At this time, developers (including companies) do often not allow to see the actual implementation of the algorithms. From a societal point of view it seems good to encourage development parties to share the algorithms used, and for researchers by providing a good example by always sharing theirs. There is also a task for editors of journals, for example by encouraging the publication of "permanent" links to such information. The group discussed several examples to indicate that in the long run taking accountability, being transparent, being credible, and showing integrity, can actually make a BCSS more effective. For one, this shift away from black-box thinking could help to explain to users *why* certain recommendations are made.

Challenge 2: Which persuasive design features are suitable for data-driven coaching? And why (not)? The group discussed which persuasive features are suitable for data-driven coaching. The main conclusion of the discussion was that 'persuasion patterns' differ according to the stage of a user. For example, the needs of a newly diagnosed diabetes patient regarding automated coaching probably differ from the needs of a more 'experienced' patient. This might require a move towards user-phase adaptive coaching based on the user profile and log data regarding the use of systems where the system provides 'active coaching' and primary task support for new users. Over time, users of a BCSS will probably become aware of their own behavior and some habits are formed. In this phase the coaching application should provide "maintenance coaching" using unobtrusive interaction and communication. When a user tends to fall back into old habits of behavior, the BCSS should go back to the "active coaching" phase. This time, the content and subjects in the active coaching phase should differ from the content and topics provided at the first use. After all, users probably know the basic skills and coaching could therefore focus on personal goals to prevent fall backs in the future. The persuasive design features are all suitable for data-drive coaching and can be selected based on the data collected during the use of a BCSS.

5 Conclusions & take home message

Based on the papers and discussion we can conclude that to benefit from the persuasive opportunities offered by data-driven BCSSs, researchers and developers need to deploy a good rationale or strategy for using extensive data for the development and evaluation of their BCSSs and report this clearly. Obviously, data quality and data analysis should be accurate for the used approach. Finally, interpretation of the data and feedback to the user requires to be provided timely and in a relevant matter. From the discussions it is apparent that there is still plenty of work to be done to address this and other points to create more effective BCSSs. This work includes the development of more transparent systems and information about the working mechanisms, as well as new opportunities from a user-phase adaptive coaching approach. Both challenges can improve our next generation of BCSSs and, provide an interesting direction for the next BCSS workshops to focus on.

Acknowledgements. We like to thank all participants for their contributions to the workshop, both their papers and the invaluable input during the discussions. Especially a big thanks go out to Markku Kekkonen, Ryan Wold, Dominic De Franco, Laurence Alpay, Piiastiina Tikka, and Aniek Lentferink.

6 References

1. Op den Akker, R., Klaassen, R., Bul, K., Kato, P., Van der Burg, G.J., Di Bitonto, P.: Let them play: experiences in the wild with a gamification and coaching system for young

diabetes patients. In: Proceedings of PervasiveHealth '17 11th EAI International Conference on Pervasive Com-puting Technologies for Healthcare. (2017)
2. Oinas-Kukkonen, H.: A foundation for the study of behavior change support systems. Personal and Ubiquitous Computing 17, 1223-1235 (2013)
3. Sieverink, F., Kelders, S., Poel, M., van Gemert-Pijnen, L.: Opening the Black Box of Electronic Health: Collecting, Analyzing, and Interpreting Log Data. JMIR Res Protoc 6, e156 (2017)
4. Michie, S., Yardley, L., West, R., Patrick, K., Greaves, F.: Developing and Evaluating Digital Interventions to Promote Behavior Change in Health and Health Care: Recommendations Resulting From an International Workshop. J Med Internet Res 19, e232 (2017)
5. Patrick, K., Hekler, E.B., Estrin, D., Mohr, D.C., Riper, H., Crane, D., Godino, J., Riley, W.T.: The Pace of Technologic Change: Implications for Digital Health Behavior Intervention Research. American Journal of Preventive Medicine 51, 816-824 (2016)
6. Klaassen, R., Van Delden, R.W., Cabrita, A.M., Tabak, M.: AIRPlay: Towards a 'Breathgiving' Approach. In: Proceedings of the Fifth International Workshop on Behavior Change Support Systems (BCSS'17) in conjunction with the 12th International Conference on Persuasive Technology, pp. 38-45. (2017)
7. Resnicow, K., Strecher, V., Couper, M., Chua, H., Little, R., Nair, V., Polk, T.A., Atienza, A.A.: Methodologic and Design Issues in Patient-Centered e-Health Research. American Journal of Preventive Medicine 38, 98-102 (2010)
8. Moller, A.C., Merchant, G., Conroy, D.E., West, R., Hekler, E., Kugler, K.C., Michie, S.: Applying and advancing behavior change theories and techniques in the context of a digital health revolution: proposals for more effectively realizing untapped potential. Journal of Behavioral Medicine 40, 85-98 (2017)
9. Hekler, E.B., Klasnja, P., Riley, W.T., Buman, M.P., Huberty, J., Rivera, D.E., Martin, C.A.: Agile science: creating useful products for behavior change in the real world. Translational Behavioral Medicine 6, 317-328 (2016)
10. Han, J.Y.: Transaction logfile analysis in health communication research: Challenges and opportunities. Patient Education and Counseling 82, 307-312 (2011)
11. Wentzel, M., Nijdam, L.: Understanding eHealth use from a Persuasive System Design perspective: an Antibiotic Information Application for Nurses. International journal on advances in life sciences 6, 210-219 (2014)
12. Jong, N.d., Wentzel, M., Kelders, S., Oinas-Kukkonen, H., Gemert-Pijnen, J.: Evaluation of Perceived Persuasiveness Constructs by Combining User Tests and Expert Assessments. In: Proceedings of the Second International Workshop on Behavior Change Support Systems co-located with the 9th International Conference on Persuasive Technology (PERSUASIVE 2014), pp. 7-15. (2014)
13. Sieverink, F., Kelders, S.M., Braakman-Jansen, L.M., van Gemert-Pijnen, J.E.: The added value of log file analyses of the use of a personal health record for patients with type 2 diabetes mellitus: preliminary results. Journal of diabetes science and technology 8, 247-255 (2014)
14. Kekkonen, M., Oinas-Kukkonen, H., Tikka, P.: Systems Supporting Occupational Health Behavior Change: A Systematic Literature Review. In: Proceedings of the Sixth International Workshop on Behavior Change Support Systems (BCSS'18) in conjunction with the 13th International Conference on Persuasive Technology, pp.
15. Manninen, I., Tikka, P.: Developing a Gamified Behavior Change Support System: Case Implicity – The Food Game. In: Proceedings of the Sixth International Workshop on

Behavior Change Support Systems (BCSS'18) in conjunction with the 13th International Conference on Persuasive Technology, pp.
16. De Frano, D., Pease A., Snaith, M.: Measuring Persuasiveness in Behaviour Change Support Systems. In: Proceedings of the Sixth International Workshop on Behavior Change Support Systems (BCSS'18) in conjunction with the 13th International Conference on Persuasive Technology, pp.
17. van Delden, R., Reidsma, D.: Automatic Measurements of a Leisure Activity for People with Profound Disabilities. In: Proceedings of the Sixth International Workshop on Behavior Change Support Systems (BCSS'18) in conjunction with the 13th International Conference on Persuasive Technology, pp.
18. Cruz-Martínez, R.R., Sieverink, F., Wesselink, R., van Gemert-Pijnen, L.: Towards Data-Driven Persuasive Coaching in a Heart Failure Telemonitoring Technology. In: Proceedings of the Sixth International Workshop on Behavior Change Support Systems (BCSS'18) in conjunction with the 13th International Conference on Persuasive Technology, pp.