Magnus Bang
Eva L. Ragnemalm (Eds.)

Persuasive Technology

Design for Health and Safety

LNCS 7284

7th International Conference, PERSUASIVE 2012
Linköping, Sweden, June 2012
Proceedings

Springer

Lecture Notes in Computer Science 7284

Commenced Publication in 1973
Founding and Former Series Editors:
Gerhard Goos, Juris Hartmanis, and Jan van Leeuwen

Magnus Bang Eva L. Ragnemalm (Eds.)

Persuasive Technology

Design for Health and Safety

7th International Conference, PERSUASIVE 2012
Linköping, Sweden, June 6-8, 2012
Proceedings

Springer

Volume Editors

Magnus Bang
Linköping University, Sweden
Department of Computer and Information Science
Linköping, Sweden
E-mail: magnus.bang@liu.se

Eva L. Ragnemalm
Linköping University, Sweden
Department of Computer and Information Science
Linköping, Sweden
E-mail: eva.ragnemalm@liu.se

ISSN 0302-9743 e-ISSN 1611-3349
ISBN 978-3-642-31036-2 e-ISBN 978-3-642-31037-9
DOI 10.1007/978-3-642-31037-9
Springer Heidelberg Dordrecht London New York

Library of Congress Control Number: 2012939234

CR Subject Classification (1998): H.4-5, J.3, K.4.1, K.4.4, I.2

LNCS Sublibrary: SL 3 – Information Systems and Application, incl. Internet/Web and HCI

Typesetting: Camera-ready by author, data conversion by Scientific Publishing Services, Chennai, India

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

Preface

Persuasive computing research is maturing. Since the First International Conference on Persuasive Technology for Human Well-being, held in Eindhoven in 2006, there have been five international conferences on the topic. Being fundamentally an interdisciplinary research field, persuasive technology integrates developments from psychological, social and cognitive research with computer science with the goal of understanding how interactive computer systems can be designed to support positive behavior change. Typical domains for persuasive technology are health, safety and education. The research area is also characterized by a healthy dose of technical curiosity. For example, the emergence of smart phones, tablets and sensor technologies as well as social networking makes it possible to create new classes of interesting and effective persuasive-technology applications. Moreover, a core of theoretical knowledge is emerging in the field as a basis for the design and evaluation of persuasive technologies.

The 7$^{\text{th}}$ International Conference on Persuasive Technology showed this multidisciplinary diversity in the contributions that spanned low-tech diaries that support cognitive behavior therapy, smart phone-based games that encourage physical activity and training, to ethical perspectives on persuasive technology. There are 21 full papers and five short papers presented in this volume. The poster session encompassed a similar amount of contributions that were published in a separate volume by Linköping University Electronic Press.

Three keynote speakers introduced current and future topics in persuasive technology; Maurits Kaptein from Eindhoven University of Technology talked about persuasion profiles and personalization. B.J. Fogg from Stanford University followed up with a talk on tiny habits and motivation waves. Finally, Harri Oinas-Kukkonen from the University of Oulu spoke about the future of persuasive computing. This year's conference also featured a set of workshops and tutorials that were held as parallel sessions. For example, HBIS 2012 – The First International Workshop on Human Behavior for Industrial Safety – featured a new exiting area for behavior change.

Persuasive 2012 was hosted by the Department of Computer and Information Science, Linköping University, Sweden. We want to express our gratitude to the reviewers and Program Committee for their invaluable comments on the contributions. We also want to thank the Swedish Energy Agency – our main sponsor – for their support as well as our contributing sponsors the Swedish National Road and Transport Research Institute, New Tools for Health, Swedish ICT and Santa Anna IT Research Institute.

June 2012

Magnus Bang
Eva L. Ragnemalm

Organization

Conference Chair

Magnus Bang	Linköping University, Sweden

Poster Chair

Eva L. Ragnemalm	Linköping University, Sweden

Organizing Committee

Magnus Bang	Linköping University
Eva Ragnemalm	Linköping University
Jonas Lundberg	Linköping University
Mattias Arvola	Linköping University
Jonas Rybing	Linköping University

Local Administration

Lise-Lotte Andersson	Linköping University

Review and Scientific Program Committee

Jack Andersen	Royal School of Library and Information Science, Denmark
Jeremy Bailenson	Stanford University, USA
Nilufar Baghaei	Unitec, New Zealand
Shlomo Berkovsky	CSIRO, Tasmania, Australia
Timothy Bickmore	Northeastern University, USA
Robert Biddle	Carleton University, USA
Winslow Burleson	Arizona State University, USA
Harry Bruce	University of Washington, USA
Samir Chatterjee	Claremont Graduate University, USA
Brian Cugelman	AlterSpark Consulting, Canada
Janet Davis	Grinell College, USA
Susan Ferebee	University of Phoenix, USA
BJ Fogg	Stanford University, USA
Jill Freyne	CSIRO, Australia
Luciano Gamberini	University of Padova, Italy
Rosanna E. Guadagno	University of Alabama, USA
Anton Gustafsson	Interactive Institute, Sweden

Jaap Ham	Eindhoven University of Technology, The Netherlands
Curtis P. Haugtvedt	Ohio State University, USA
Jette Hyldegaard	Royal School of Library and Information Science, Denmark
Stephen Intille	MIT and Northeastern University, USA
John Ittelson	California State University, USA
M. Sriram Iyengar	The University of Texas, Houston, USA
Giulio Jacucci	Helsinki Institute for Information Technology, Finland
Rilla Khaled	IT University of Copenhagen, Denmark
Judith Masthoff	University of Aberdeen, UK
Tom MacTavish	Illinois Institute of Technology, USA
Cees Midden	Eindhoven University of Technology, The Netherlands
Fred Muench	CASA, Columbia University, USA
Harri Oinas-Kukkonen	University of Oulu, Finland
Thomas Ploug	Copenhagen Institute of Technology, Denmark
Wolfgang Reitberger	Vienna University of Technology, Austria
Martha Russell	Stanford University, USA
Timo Saari	Tampere University of Technology, Finland
Anna Spagnolli	University of Padova, Italy
Jason Tester	Institute for the Future, USA
Julita Vassileva	University of Saskatchewan, Canada
Peter de Vries	University of Twente, The Netherlands
Fahri Yetim	University of Siegen, Germany
Johan Åberg	Linköping University, Sweden
Peter Øhrstrøm	Aalborg University, Denmark

Full Sponsor

The Swedish Energy Agency

Contributing Sponsors

Swedish National Road and Transport Research Institute, New Tools for Health, Swedish ICT and Santa Anna IT Research Institute.

Table of Contents

Determining the Determinants of Health Behaviour Change through an Online Social Network

Noreen Kamal and Sidney Fels

Electrical and Computer Engineering, University of British Columbia,
Vancouver, BC
{noreenk,ssfels}@ece.ubc.ca

Abstract. The ABC framework provides determinants for leveraging the motivational power of online social networks with the determinants for promoting health behaviour changes. We designed VivoSpace, a medium fidelity prototype of an online social network to promote healthy behaviour changes based on the guidelines for incorporating these determinants. We evaluated the determinants of appeal, belonging and commitment using both direct and indirect methods with 36 adult subjects. Indirect evaluation methods included a helping game experiment, adopted from experimental behavioural economics to measure indirect reciprocity evoked by VivoSpace, which is an important factor in developing belonging. Similarly, an in-group experiment was adopted to evaluate group commitment. Our results show that VivoSpace's design based on the ABC framework result in a strong degree of agreement with the appeal determinants with evidence for the promotion of belonging and commitment. Thus, we have evidence for the effectiveness of design elements for evoking behaviour change to improve health using an online social network.

Keywords. Laboratory experiments, health behaviour change, VivoSpace, ABC framework.

1 Introduction

Designing technologies to promote health behaviour change has become an area of interest both in research and for commercial products. Various technologies that promote health behaviour change have been developed and evaluated. These systems include: *Houston* a mobile phone application that shares step counts with friends [4]; *UbiFit* a mobile phone application that provides a garden display for one's exercise performance over time [5,6]; and *MAHI* a mobile phone application for diabetics that allows entry of experiences and reflection with health educators [14]. This is complimented by an increased interest in the use of online social networks for health behaviour change [15], which is not surprising as one's social networks greatly influence health behaviour [7]. Furthermore, there are numerous applications available commercially to assist users in leading healthy lifestyles. These commercial applications are available in many modalities including: online applications[1], online social

[1] Examples include SparkPeople (www.sparkpeople.com) & DailyBurn (www.dailyburn.com).

M. Bang and E.L. Ragnemalm (Eds.): PERSUASIVE 2012, LNCS 7284, pp. 1–12, 2012.

network games[2], console games[3] and pervasive/wearable systems[4]. Many of these commercial applications include social aspects to provide social support to users in making health behaviour change. Few attempts have been made to articulate design principles by harnessing the popularity of online social networks to encourage healthier living based on theoretical results from combining health and online social network literature. One such framework, the ABC (Appeal, Belonging, Commitment) Framework, has been proposed that establishes a set of determinants that can be addressed in an online social network designed for changing health behaviour [11].

In this paper, we investigate and measure whether the determinants of the ABC framework are met in an online social network designed to motivate health behaviour change. For our investigation, we have developed VivoSpace an online social network, based on the design principles that can be extrapolated from the determinants of the ABC framework. We then evaluate VivoSpace in a laboratory setting with the aim to determine if the online social network system can motivate health behaviour change, based on the determinants of the ABC framework. The determinants in the Appeal dimension can be evaluated through direct inquiry; however, challenges surface when evaluating Belonging and Commitment. Therefore, we use two indirect methods: 1) a helping game is employed to evaluate belonging; and 2) a group performance exercise is used to evaluate commitment.

The key contribution of this paper is the use of the ABC Framework to evaluate the design of an online social network, VivoSpace. Furthermore, this paper uses triangulation of direct and two indirect methods in evaluating an online application's ability to influence human behaviour.

2 Related Work

There are several persuasive technologies that have been designed and evaluated to change health behaviour. The *Houston* system is a mobile phone application that tracks and shares progress towards a step count goal with a group of friends [4]. This study evaluated the determinants of *goal setting* and *sharing information*. They found that over a short period of time those that shared data met their goals more than those that did not share their data. Both *goal setting* and *social factors* are considered in the ABC framework, as being some of the determinants of behaviour change; however, there are many other determinants, including *self-efficacy*, *knowledge* and *attitude toward the behaviour* that also need to be considered.

UbiFit is another mobile phone application that has been designed and evaluated [5,6]. The *UbiFit* was initially designed based on the *Goal-Setting Theory*, and evaluated based on where the goal comes from: the user, guidelines, medical/health expert and group set [5], and these were evaluated qualitatively. The user set and expert set conditions were the most popular. This study employed the *goals* and *knowledge*

[2] An Example is the facebook application HealthSeeker (`healthseekergame.org`).
[3] Examples include Wii Fit (`wiifit.com`) and EA Sports Active (`www.easportsactiveonline.com`).
[4] Examples include fitbit (`www.fitbit.com`) and NikePlus (`nikeplus.com`).

determinants, which are important in the ABC framework, but there are several other determinants, such as *self-efficacy, attitude towards the behaviour*, and *social norms* that also need to be considered to truly change health behaviour. The *UbiFit* system was also designed based on the Goal-Setting Theory, Transtheoretical Model, Presentation of Self in Everyday Life and Cognitive Dissonance Theory to develop design principles, which were then evaluated [6]. Although, the determinants from these theories were not drawn out in the work, the use of several theories to design and evaluate the system provides greater depth in understanding the motivations for changing health behaviour.

The *MAHI* system is a mobile phone application that was developed for diabetics to monitor their health and reflect upon it through social interaction with diabetes educators [14]. The *MAHI* system was used on individuals that have lived with diabetes for a long time [14]. The qualitative evaluation was based on the textual entries made by users of *MAHI*; they found that the system provided a forum for "constructing identities". Although not explicit in this study, this study does reveal the *shared identities* determinant for using social systems, which is defined by the ABC framework.

3 ABC Framework

Existing theoretical models provide the determinants for the motivations for using online social networks and the motivations for changing health behaviour. The theoretical models for use of online social networks include: the Uses and Gratification Theory [12], Common Identity and Common Bond Theory [18], Social Identity Theory [8], and the Theory of Organizational Commitment [2,12]. The theoretical models for motivating health behaviour change include: the Health Belief Model [10], Social Cognitive Theory [3], Theory of Planned Behaviour [1], the Common Sense Model [13], and the Transtheoretical Model [17]. Extracting the determinants for behaviour change from these theoretical models reveal three dimensions for online social networks that motivate health behaviour change: appeal (individually based), belonging (socially based) and commitment (temporally based) [11]. Together these dimensions provide the foundation for the Appeal Belonging Commitment (ABC) Framework illustrated as a text cloud in Figure 1 that fully describes how online social networks can be used to motivate health behaviour change. The figure shows that health behaviour change and use of online social networks are complex and are defined by a multitude of factors that have significant interplay.

Since the ABC framework is based on a review of existing theoretical models, we argue that the application of the ABC framework in the design and evaluation of online social networks will reveal a system that will have a high likelihood of influencing long-term behaviour change. Therefore, we are evaluating the design of VivoSpace based upon the determinants of the ABC framework. Based on this framework, online social network first needs to Appeal to users on an individual level in order for

the system to be used. The determinants for use then promote the determinants that ensure that the new health behaviours Appeal to users. Once, the online social network is appealing, it also needs to promote Belonging, which then promotes social norms to alter health behaviour. Finally, Commitment needs to occur to ensure habitual use of the online social network, and maintenance of newly acquired health behaviour also needs to occur.

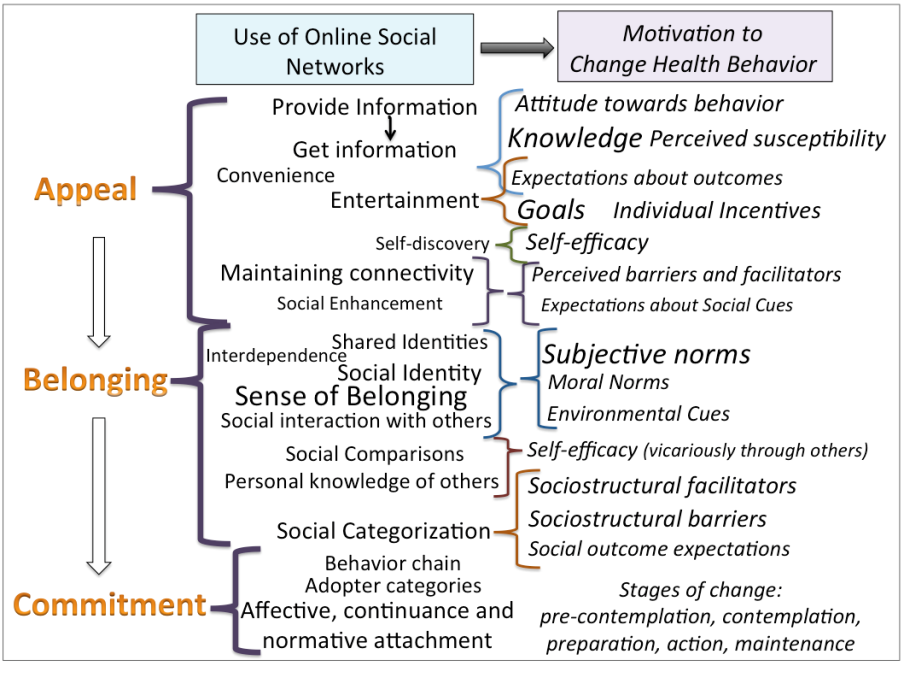


Fig. 1. Appeal Belonging Commitment Framework for using online social networks for health behaviour change showing the major (large text) and minor (small text) determinants for behaviour change and the interplay between the determinants (brackets)

4 The VivoSpace Health Social Network

We have developed a medium fidelity prototype called VivoSpace, which is an online social network that is designed to motivate health behaviour change. The design strategies were drawn from the determinants of the ABC framework. The main activity page for VivoSpace's medium fidelity prototype is shown in Figure 2. This medium fidelity prototype was developed using HTML, CSS, Javascript and jQuery with the vision to present a realistic interactive representation of our system. In total there were 32 HTML pages, 1 css file, and 2 javascript files.

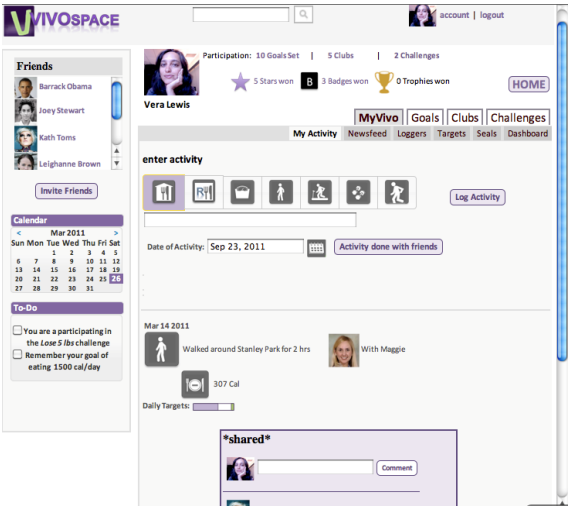


Fig. 2. A screenshot of the VivoSpace prototype, this page is the main activity page

There are 6 main functions that the medium fidelity prototype provides:

1. **Entry of meals and physical activity.** Provides a means to enter meals or activity (Figure 2); cumulative daily values for each of the nutrients are displayed with any evidence-based seals if it meets set criteria. The user can change the types of activities that s/he may log, the nutrients that are displayed, and the daily targets.
2. **Newsfeed.** The newsfeed page shows activities and accomplishments that friends shared. Any logged activity above can be shared with one's social network, and then it will appear on their friends' newsfeed page, and they can comment on it.
3. **Dashboard.** The dashboard page shows a summary of how the user is doing in terms of the nutritional intake based on the daily targets that are set. Time series charts are also provided along with a green checkmark for those nutrients where the target is met, and a red "*high*" or "*low*" where the target is not met.
4. **Personal goals.** The goals section provides the ability to track and create new health goals, and to view rewards for successes. The current active goals are shown with the nutritional intake compared against the target goals; charts and definitions of nutrients are also provided. New goals can also be created from a library of goals such as lose weight and healthy heart. Stars are earned when goals are successfully completed; users can view their own stars and their friends' stars.
5. **Group goals (clubs).** Clubs are similar to goals with the exception that user's friends are invited to participate in the club. When viewing current clubs users can see how other members of the club are doing. Members of the club can comment on people's progress in the club. Badges are earned when members successfully complete the requirements. Users can view and comment on their friend's badges.
6. **Competitive goals (challenges).** Challenges are similar to clubs with the exception that there are competitive and there is only one winner. Challenges display a leaderboard. The winner of challenges earns a trophy.

5 Experiments and Results

VivoSpace was evaluated based the determinants of the ABC framework, which involved using multiple laboratory experiments. The purpose was to determine if the design of VivoSpace provided the determinants for use of the online social network and motivation to change health behaviour. A total of 36 adults participated in the experiments (described in Section 5.1). They were recruited through university listserves, posters located around campus and through advertisements on Craig's list.

The determinants of appeal were measured directly through questionnaire feedback after the participant completed specific tasks on the medium fidelity prototype. Some of the *belonging* determinants were also enquired directly, and these methods are explained in detail in Section 5.2. However, measuring *belonging* cannot truly be measured through direct inquiry, as self reports about belonging are unreliable as actual sense-of-belonging are often divergent from how an individual perceives their belonging to a group. For this reason, we turned to the field of behavioural economics and adopted the *helping game experiment* [19] to evaluate if VivoSpace evokes indirect reciprocity. The methods and results for how this experiment was adapted for the evaluation of the medium fidelity prototype is described in Section 5.3. Similarly, *Commitment* is a temporal dimension and difficult to measure in a 1-2 hour experiment; however, the *in-group experiment* [9] is an indirect means to evaluate if VivoSpace promotes group commitment (described in Section 5.4).

18 of the 36 participants (Participant A) completed all three experiments. These 18 participants first complete the individual task experiment, which takes approximately 1 hour to complete. Participant B (18 total) then joins Participant A, and they complete the helping game and group commitment experiments together in pairs, which takes another 1 hour to complete. Each participants in the A group were remunerated $10 and each participant in the B group were remunerated $5.

5.1 Participants

In the Participant A group, 10 were male and 8 were female. There were 7 aged 19-24, 5 aged 25-34, 5 aged 35-49 and 1 was aged 50-64. The distribution for self-reported ethnic identity for this group was as follows: 10 were Canadian, 3 were West Asian, 1 was Chinese, 1 was Hispanic, 1 was First Nations, 1 was European and 1 was Australian. The majority were students: 5 were undergraduate students and another 5 were graduate students. Of the remaining participants, 4 were unemployed, 1 was a canvasser, 1 was a cleaner, 1 was a postdoctoral fellow and 1 was a physician.

In the Participant B group, 10 were male and 8 were female. There were 12 aged 25-34 and 6 aged 35-49. 8 were Canadian, 3 were West Asian, 3 were European, 1 was South Asian, 1 was Australian, 1 was Chinese and 1 was American. 6 were graduate students and 3 were software developers. The occupations of the remaining 9 were: employment assistance worker, postdoctoral fellow, professor, college teacher, researcher, software development manager, scientist, education program manager, and administrator.

This demographic distribution of participants shows that a good degree of gender, age, ethnic and occupational distribution was achieved. This purpose of achieving diversity is to provide an understanding of how a broad population would be motivated to use VivoSpace and through its use be motivated to change their health behaviour.

5.2 Individual Task Experiment

The individual task experiment methodology is drawn from traditional usability tests. During this experiment, each participant is asked to complete a group of tasks on VivoSpace. There are six groups of tasks for each of the prototype's functions described in Section 4. After each task group, the participant is given a questionnaire to complete, which contains statements that correlates to determinants from the ABC framework. The participants provide their level of agreement or disagreement to each statement using the 7-point likert scale, where 1 is strongly disagree and 7 is strongly agree. Most of the questions enquired about determinants from the appeal dimension, but there are also some from the belonging dimension.

Examples of the questions and the corresponding determinant from the ABC framework are the following:

"I would be able to gain information about myself and my capabilities by using a system like this" (Appeal: self-discovery)

"The newsfeed would allow me to view how my friends and family are staying healthy" (Belonging: social comparison)

The results of the 7-point likert questionnaire provide an understanding of how each of the 6 task groups promoted the appeal determinants and some belonging determinants of behaviour being sought through VivoSpace.

The results for the first task group (entry of meals and physical activity) show that the entering activity task group most strongly endorses the following determinants (mean > 5.0): *self-discovery*, *get information*, *convenience* and *health outcomes*. The design does not favor the determinant *to provide information* (mean=3.67). The remaining determinants, *entertainment*, *social enhancement*, and *environmental cues*, are neither strong nor weak (4.0<mean<5.0). Factorial ANOVA analysis on these three determinants was performed, where the factors were gender and age group. There was a statistically significant difference between gender for finding this task group *entertaining* ($F(1,12)=4.6$, $p=0.05$), where the mean difference between male and female is -1.4 with females finding it more entertaining then men. The results from second task group (newsfeed) show very good agreement to both the *appeal* and *belonging* determinants with a mean for all responses of over 5.0. The results from third task group's inquiry (dashboard) shows strong agreement for many of the *appeal* determinants especially *get information*, *self-discovery*, *convenience* and *knowledge* (mean>=5.94). There is less agreement in the dashboard providing any value to *perceived barriers & facilitators* and it also does not provide a very strong *incentive* to live healthier (mean < 5). The results from fourth task group (goals) show good agreement with the appeal determinants, but it was not very *entertaining* (mean=4.89), and they would not use it to *enhance their social position* (mean=4.11).

Figure 3 shows results for the fifth (clubs) and sixth (challenges) task group. There was good agreement that the design of the clubs' pages with the appeal and belonging determinants, but the design of the challenges' pages scored lower for almost all determinants. Further, the design did not support *social enhancement* especially for clubs. Further analysis factorial ANOVA found that there is a statistically significant difference in the challenges task group between male and female in *social enhancement* ($F(1,12)=5.777$, $p=0.033$) with a mean difference between male and female of 1.8. Statistically significant difference in the interaction between gender and age groups for *subjective norms* ($F(2,12)=4.22$, $p=0.041$) was found showing that different genders in different age groups felt that challenges promoted *subjective norms*.

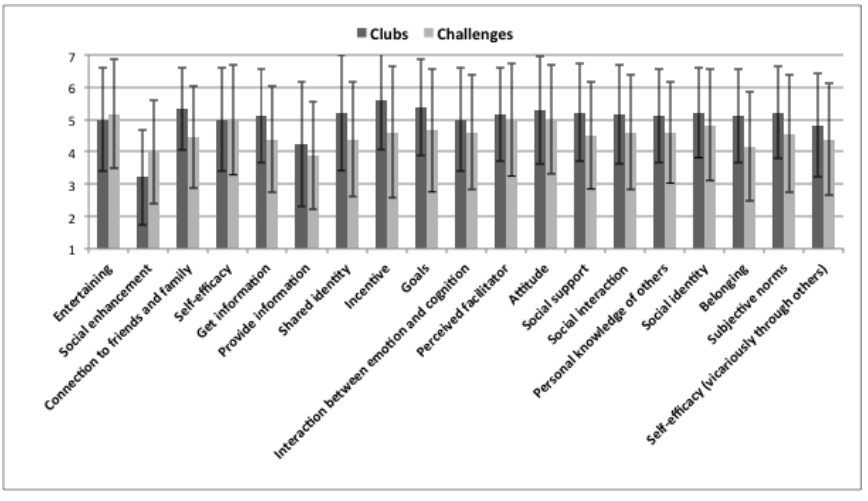


Fig. 3. Appeal and belonging determinant 7-point likert responses for the *Clubs* and *Challenges* task group

5.3 Helping Game Experiment

The determinants in the belonging dimension of the ABC framework are difficult to measure through direct inquiry methods such as those described above in the individual task experiment. Therefore, we utilize indirect experimental methods to better understand if our design contributes to belonging and triangulate the findings from the experiment above. Indirect reciprocity is a critical factor that facilitates belonging, and other related determinants such as social categorization and group comparison. The Helping Game Experiment originates from experimental behavioural economics has been shown to provide an understanding of indirect reciprocity and group belonging [19]. This experiment is based on the "repeated helping game" developed by Nowak and Sigmund [16]. This experiment evaluates indirect reciprocity, which is necessary for online social networks to develop social interaction and cooperation through group participation, feedback and discussion. We adopted the experiment described by Seinen et al. [19] to be applied to the evaluation of VivoSpace in a laboratory setting. In the Seinen et al. experiment, there was no actual activity that the

randomly assigned pairs were helping each other on. We adapt this experiment to be applied to providing help in completing tasks on VivoSpace.

In our experiment, after Participant A has completed the individual task experiment, s/he is joined by Participant B. A and B do not know each other. Participant A is now an expert on VivoSpace and understands how to use it, and Participant B is a novice. This set-up provides a good backdrop to apply the Helping Game Experiment, as Participant B must now complete the tasks and Participant A decides to help or not. Prior to starting a task, Participant A must decide if s/he will help Participant B at a cost of $1 or $5 depending on the cost of the task. Every time A helps B, s/he earns a point. There are two conditions: in the first condition, A is told that B already has 17 points; in the second condition, no information is provided on B's points.

The mixed design experiment has 2 variables with two conditions in each variable. Therefore, the 4 conditions in this mixed design are: 1) information provided and high cost, 2) information provided and low cost, 3) no information provided and high cost, and 4) no information provided and low cost. We hypothesize that the effect of helping on VivoSpace will produce different results from the Seinen et al. [25] experiment, which found that having information about their partner's points resulted in a statistically significant difference in providing help.

A mixed design factorial 2x2 ANOVA is run on the results of the Helping Game Experiment. The dependent variable is the percent of times Participant A helped Participant B. The within subject variable is the cost (cheap or expensive) for the tasks. The between subject variable is information provided or no information provided about Participant B's score. The results show that there is a significant difference between the cheap and expensive tasks ($F(1, 16)=5.54$, $p=0.032$). The mean difference between the cheap and expensive is 23.86%. This is not surprising and consistent with the Seinen et al. experiment [19], which shows that there is validity to the use of fake money as the cost of the tasks played a role in participant's decision to help or not help. The results for having information of B's score in the 2x2 factorial ANOVA reveals divergent results from the Seinen et al. experiment [19]. Interestingly, there is no statistically significant difference between the two conditions: information and no information ($F(1, 16)=0.51$, $p=0.386$). The top graph in Figure 4 shows the estimated marginal means and standard deviation for the percent of tasks that Participant A helped in the mixed design helping game experiment.

5.4 Group Commitment Experiment

Although measuring commitment to VivoSpace and new health behaviour is not possible in a laboratory experiment, group commitment has been measured in the laboratory by social psychologists. If VivoSpace can invoke group commitment, it will assist in developing commitment to the VivoSpace system and commitment to new health behaviours. The group commitment experiment has been adapted from Ellemers et al. experiment [9]. The premise of the experiment is that group commitment can be measured by creating groups, where the group boundaries are permeable, and a situation is imposed where group status is provided and compared to other groups.

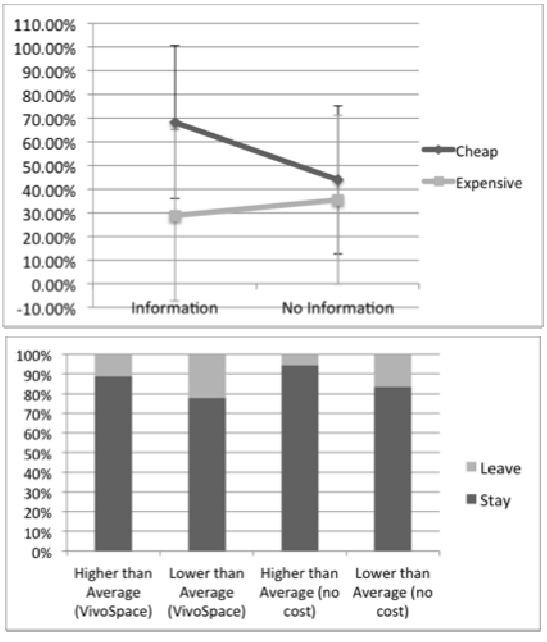


Fig. 3. Top – Helping game experiment's estimated marginal mean values and standard deviation for percent of tasks helped (n=9 per condition). Bottom – Group commitment experiment's percent willing to stay or leave with their partner (n=18 per condition)

Again, we use a mixed design experiment. The same partners from the helping game experiment are now told that they are team, and must complete a 10-question multiple-choice test on VivoSpace, where questions are not readily evident in the VivoSpace prototype. After the test, the participants are separated and provided with their score and the overall average score for this test. One person is told they scored 70% and the average for the test is 50%, and the other person is told they scored 70% and the average for the test is 90%. The two conditions are counter balanced between participants A and B. Each participant is then given the option of group mobility with two separate hypothetical caveats (within subjects): 1) they will not be able to use VivoSpace, and 2) it would cost them $50 to leave. This mixed design experiment has four conditions: 1) above average stay with VivoSpace, 2) above average stay at no cost, 3) below average stay with VivoSpace, and 4) below average stay at no cost.

The results of the group commitment experiment reveal that participants were inclined to stay with their groups rather than leave. The results are shown on the bottom graph of Figure 4. Statistical analysis was run on this 2x2 mixed design through Chi-Square test of association. The results show no statistically significant results between the test performance condition (higher and lower than average) and the choice to stay for VivoSpace (Chi-Square(1)=1.639, p=0.200). Furthermore, there was no statistical difference between the test performance and the choice to stay with the group at no financial cost (Chi-Square(1)=1.172, p=0.279).

6 Discussion

The ABC framework provided a foundation to measure the determinants for use of an online social network and its ability to lead to health behaviour change. When the determinants of *appeal* were enquired through questionnaire responses after interacting with the system, it was found that many of the determinants were met through the design of VivoSpace. However, there was variation in the responses between the clubs and challenges components of the design especially between male and females. For this reason the design can be iterated to combine goal, clubs and challenges into a single component that allows for participants to decide to invite friends and make it competitive or collaborative. The *individual task experiment* also revealed the weakest area of the design: motivation to *provide information*, which is an integral part of the system. Without data being logged, the other components in the design become less meaningful as only partial health information may be logged. Therefore, the design should be iterated to ensure stronger motivation to provide information.

The *helping game experiment* provide evidence that the VivoSpace prototype promotes indirect reciprocity and by extension group belonging. Since there was no statistically significant difference between having and not having information about their partner's score (status), we can assume that VivoSpace played a part, which is because in the Seinen et al. experiment, knowledge of their partner's score did play a statistically significant difference than those who did not have this information [19]. Although there are many variables that could interfere with this experiment, these are encouraging results. By showing that there was no statistically significant difference, the motivation to help or not help would be based on the design of the prototype rather than their partner's score.

The *group commitment experiment* shows that VivoSpace can provide group commitment. Participants were willing to stay with their groups even those participants that performed worse than average. Furthermore, when the cost to leave the group was the loss of use of VivoSpace or financial, there was still no statistically significant difference. Showing that VivoSpace does promote commitment as much as economic gain. Although the answers to these questions were made privately, there is a risk that participants are not willing to admit their desire to leave the group.

7 Conclusion

A medium fidelity prototype for VivoSpace was evaluated based on the ABC Framework. We measured *appeal* and *belonging* through an individual task experiment, where each participant was asked about components of the design based on appeal and belonging determinants. It was found that VivoSpace did meet most of the appeal determinants; however, the design could be improved to ensure that all users would find the components of VivoSpace appealing. Furthermore, the motivation to provide personal information was not as appealing as required in order for VivoSpace to be useful. The design of VivoSpace does promote indirect reciprocity, which can lead to group belonging, as can be interpreted through the *helping game experiment*. Finally,

the design of VivoSpace does promote group commitment, as participants of the *group commitment experiment* were willing to stay with their partners despite half the participants learning of poor scores.

References

1. Ajzen, I.: The Theory of Planned Behaviour. Organ. Behav. Hum. Dec. 50, 179–211 (1991)
2. Allen, N.J., Meyer, J.P.: The measurement and antecedents of affective, continuance and normative commitment to the organization. J. Occup. Psychol. 63, 1–18 (1990)
3. Bandura, A.: Health promotion through social cognitive means. Health & Edu. Behaviour 31(2), 143–164 (2004)
4. Consolvo, S., Everitt, K., Smith, I., Landay, J.A.: Design requirements for technologies that encourage physical activity. In: CHI 2006, pp. 457–466 (2006)
5. Consolvo, S., Klasnja, P., McDonald, D.W., Landay, J.A.: Goal-setting considerations for persuasive technologies that encourage physical activity. In: Persuasive 2009, pp. 8:1–8:8 (2009)
6. Consolvo, S., McDonald, D.W., Landay, J.A.: Theory-driven design strategies for technologies that support behaviour change in everyday life. In: CHI 2009, pp. 405–414 (2009)
7. Christakis, N.A., Fowler, J.H.: The spread of obesity in a large social network over 32 years. New Engl. J. Med. 357(4), 370–379 (2007)
8. Dholakia, U.M., Bagozzi, R.P., Pearo, L.K.: A social influence model of consumer participation in network- and small-group-based virtual communities. Int. J. Res. Mark 21, 241–263 (2004)
9. Ellemers, N., Spears, R., Doosje, B.: Sticking together or falling apart: in-group identification as a psychological determinant of group commitment versus individual mobility. J. Pers. Soc. Psychol. 72(3), 617–626 (1997)
10. Janz, N.K., Becker, M.H.: The Health Belief Model: A Decade Later. Health Educ. Behav. 11(1), 1–47 (1984)
11. Kamal, N., Fels, S., Ho, K.: Online social networks for personal informatics to promote positive health behaviour. In: Proc. WSM 2010, pp. 47–52 (2010)
12. Lampe, C., Wash, R., Velasquez, A., Ozkaya, E.: Motivations to participate in online communities. In: CHI 2010, pp. 1927–1936 (2010)
13. Leventhal, H., Diefenbach, M., Leventhal, E.A.: Illness cognition: Using common sense to understand treatment adherence and affect cognition interactions. Cognitive Ther. Res. 16, 143–163 (1992)
14. Mamykina, L., Miller, A.D., Mynatt, E.D., Greenblatt, D.: Constructing identities through storytelling in diabetes management. In: CHI 2010, pp. 1203–1212 (2010)
15. Morris, M.E., Consolvo, S., Munson, S., Patrick, K., Tsai, J., Kramer, A.D.I.: Facebook for health: opportunities and challenges for driving behaviour change. In: Ext. Abstracts CHI 2011, pp. 443–446 (2011)
16. Nowak, M.A., Sigmund, K.: Evolution of indirect reciprocity by image scoring. Nature 393, 573–577 (1998)
17. Prochaska, J.O., Velicer, W.F.: The transtheoretical model of health behaviour change. Am. J. Health Promot. 12(1), 38–48 (1997)
18. Ren, Y., Kraut, R., Kiesler, S.: Applying common identity and bond theory to design of online communities. Organ. Stud. 28(3), 377–408 (2007)
19. Seinen, I., Schram, A.: Social status and group norms: indirect reciprocity in a repeated helping experiment. European Economic Review 50, 581–602 (2006)

Empowering Independent Living for People with Autism: Designing Supportive, Low-Cost, Interactive E-Health Environments

Winslow Burleson, Naomi Newman, and Ryan Brotman

School of Computing, Informatics, and Decision Systems Engineering,
Arizona State University
{Winslow Burleson,ncnewma1@asu.edu,Ryan Brotman}@asu.edu

Abstract. An investigation of the Caregiver Autism Residential E-health (CARE) system, a low-cost, end-user deployable smart home technology, has been evaluated for its potential as an empowering assistive technology for adults living with autism. It allows adults living with autism and their caregivers to create personalized smart home interventions that provide motivational support for activities of daily living, social relationships, and safer behaviors. This is achieved through the use of a ubiquitous computing system composed of off-the-shelf consumer electronic technologies. The Environmental Rating Scale (ERS), designed to assess residential interventions for people living with autism, guided the development and evaluation of CARE/ERS heuristics and interaction scenarios. The contributions of this investigation, advanced through an iterative design process involving expert reviewers, caregivers, and end-users in a patient centered approach for the design of actualizing e-health interventions, can be readily applied to a broad range of residential circumstances that improve quality of life.

Keywords: Autism, Smart Homes, E-Health, Multimedia Services, and Assistive Technology.

1 Introduction

Adults living with autism require significant levels of assistance, which is often very expensive and fails to adequately meet the personal needs of these individuals [2, 3, 14]. The 2010 MetLife Market Survey of Nursing Home, Assisted Living, Adult Day Services, and Home Care Costs [30] found that the national average monthly cost of an in-home hired caretaker is over $2000. The past decade's sharp increase in the population of people living with autism [2, 20] demands urgent and effective patient-centered approaches to the design and evaluation of residential interventions that support increased quality of life for people living with autism.

This investigation advanced a participatory approach for creating and evaluating residential multi-media e-health interventions for adults living with autism using the CARE system. Using visual scenario-based design activities in conjunction with established instruments, commonly used in the design of caregiver support for this

M. Bang and E.L. Ragnemalm (Eds.): PERSUASIVE 2012, LNCS 7284, pp. 13–30, 2012.

population, four case studies present the design and development of pervasive technologies that deliver services to adults with autism and other cognitive disabilities. The ERS, designed to assess residential interventions for people living with autism, guided the development and evaluation of a set of CARE/ERS heuristics that are broadly applicable to the needs of adults living with autism, others living with cognitive impairments, and their caregivers.

The CARE system couples low-cost, off-the-shelf consumer electronics with heuristics adapted from the ERS [38]. Development of CARE/ERS heuristics was also informed by heuristics for assistive robotics [36] and heuristics for the design of products for people with disabilities [39]. CARE is based on a National Science Foundation and Alzheimer Association sponsored system, called Game as Life – Life as Game (GALLAG). The core of the CARE architecture is an open framework designed to allow for the incorporation of a wide range of supportive technologies. The system currently integrates iPhones, iPod Touch devices, motion sensors, magnetic door sensors, and wearable devices (e.g., accelerometers) connecting with Perceptive Automation's Indigo 4.0 smart home server (see Figure 1) [17] and custom software modules to communicate between physical and digital environments. The server's communication between sensors and output devices allows for users' behavioral actions to trigger the launching of sound files, video, and turn on and off lights and appliances. This supportive activity-based intervention system empowers individuals and/or caregivers to co-author sets of action triggers that are personally tailored to their individual needs and sensitivities [4]. These action sets drive the design of scenarios [10] that facilitate care-oriented interventions. CARE's ubiquitous in-home support expands the nature of multimedia services and technologies for e-health as it seeks to improve quality of life in the residential environments of adults living with autism.

The background sections of this paper review the success and limitations of the existing approaches to smart home development and the evolving strategies and technologies that support people living with autism. The CARE system architecture is then presented along with case studies that informed the adaptation of the widely used and well-validated ERS [38] into a set of ERS/CARE heuristics for adults living with autism.

The CARE system and the scenario-based co-design approach to the development of heuristics for multimedia services and creation of effective technologies for e-health make contributions to the advancement of smart home assistive technologies that support independence and quality of life, not only for people living with autism, but for a broad population in need of personally tailored adaptive support.

2 Background

This section provides a review of related literature on independent residential environments that accommodate people with autism and the strategies that caregivers and providers of existing assistive technologies, including assistive smart homes, have employed.

2.1 Living Independently with Autism

Over 80% of people living with autism are under the age of 22 [2], and many will soon be entering the "launching period," defined as the phase in a person's life during which s/he prepares to live independently from his/her parents for the first time [21]. Though residential options for people living with autism or a related Autism Spectrum Disorder (ASD) have increased in both quantity and quality over the last decade [20], there is tremendous need for flexible and adaptable alternatives that allow the individual optimal amounts of both freedom and safety. With approximately 730,000 people, in the US alone, living with autism under the age of 21 [5] who may be preparing to make this transition within the next two decades, exploring quality residential options that adapt to the needs of individuals is imperative.

Ahrentzen and Steele found that 81% of adults living with autism between the ages of 19 and 30 live with a parent or guardian, and only 14% live in a supported residence for individuals with special needs [1]. Jabrink and Knapp reported that over 90% of the financial expense for a caretaker of an individual living with autism is spent paying for adult services, such as supported living [18], which is a service that approximately two-thirds of adults with autism in their 30's utilize [20]. Though the financial costs of current residential options for people living with autism are high, the benefits of living away from the family home are well documented. People living with autism who live independently are more responsible, confident and independent, and their parents are less stressed [20], than those who reside in more dependent living situations.

Over the last two decades, there has been a sharp drop in the number of people with autism living in institutions and an increase of people with autism living outside of the family home and maintaining employment [16]. However, one third of people living with autism reside with their primary caregivers [20], and 25% of these primary caregivers are over the age of 60 [15], which prompts concerns about aging parents' ability to care for adults living with autism. Given the large number of adults living with autism, the significant personal and societal benefits of independence [37], and the substantial costs of existing residential options, it is essential to advance low-cost, adaptable solutions that can be effectively scaled to better serve people living with autism and their individual needs.

2.2 Smart Homes as Residential Options

Smart homes have the potential to provide significant improvements to independent residential environments for people with autism, especially as they become: increasingly cost effective; highly adaptable; easy to modify, e.g., tailored to the end-user needs of individuals; and enhanced with a variety of specific opportunities for behavioral monitoring, modification and reward. There are several Smart home systems on the market, only a few of which are being adapted for people living with autism. A good example of a prototype Smart home that has been developed with autism in mind is The Georgia Tech Aware Home [19]. This system includes many elements targeted to people with autism. One of its features is a "Technology Coach," which

helps explain how to use home medical technology by providing feedback based on observations of the user's interactions with the devices. Oliver *et al*. developed the Ambient Kitchen project [33], a high-fidelity simulation environment consisting of a full-scale kitchen networked with sensing and feedback technologies. The space provides real-time task support of kitchen activities through audiovisual instruction and monitors task progression through sensor-enhanced kitchen tools.

Beyond technologies targeted specifically for autism, smart homes have the unique ability to efficiently aid users in a wide variety of personal tasks without the reduction in level of privacy that a caregiver presents. MavHome [11], (short for Managing An intelligent, Versatile HOME) is a low-cost Smart home that aims to increase the productivity and comfort of the user, as well as efficiency and organization of the house. Rather than simply responding to users' triggers, the MavHome is able to make decisions based on information that it detects in its environment. It can choose to turn off the sprinkler system if there is a high chance of rain and performs several common smart home tasks, such as: monitoring inventory and replacing it if it is low (e.g. ordering more milk if there is none in the refrigerator), controlling home systems (e.g. temperature, sprinklers, ventilation), and adjusting the home based on the user's previous activity (e.g. altering shower temperature to match the user's previous usage). Another exemplar of supportive smart home technology has been developed for people with dementia. AT EASE (Automated Technology for Elder Assessment, Safety, and Environmental monitoring) [26] collects, reports, and manages activity data of residents through motion sensors, flooding and overflow sensors, contact sensors, pressure sensors, and appliance on/off sensors. It implements stratified reporting of data concerning a person's health, safety, and well-being.

2.3 Technology for Living with Autism

The value of supportive interactions with technology for people living with autism and other cognitive disabilities is well documented [22, 27, 29, 30]. For example, virtual social skills training [31] using a system that integrates useful real-life scenarios, including choosing a bus seat or introducing oneself, into a virtual setting (i.e. an immersive simulator), allowing the user to receive immediate, quantitative feedback, has been shown to be beneficial. Leroy and De Leo [23] developed an application for smartphones that allows users with autism to accompany their text messages with drawings and images in order to improve communication between themselves and their caregivers. Madsen et al. [24, 35] have developed affective computing technology, which aids people living with autism in understanding social cues and facial expressions, through advance image recognition and appropriate feedback tools. Social networking sites are also valuable technology for practicing social skills. Carmien *et al.* conducted a series of studies resulting in three mobile device applications to assist adults living with cognitive disabilities, including autism, in their use of mass transit systems [9]. These applications were a personal travel assistant that interfaced directly with disabled individuals, a path planning tool for members of the disabled individual's immediate support system to program travel routes on their behalf, and a real-time tracking system to notify support system members if the individual has deviated

from the travel route. As a means of circumventing the need for nonverbal social cues and to minimize apprehension or confusion that they often experience in face-to-face interaction, people living with autism find social networking websites useful and enjoyable to communicate with people they already know and to meet new people [7].

2.4 Assessing Supportive Environments

Several heuristics for assistive environments have been developed. Those proposed by Tsui [36] for the use of supportive robotic interactions and environments assert the importance of supporting safety, trust, flexibility, and ease. They also highlight the goals of eliminating errors, reducing processing time, and minimizing complexity. Vanderheiden [39] proposed design guidelines to increase accessibility of products for persons with disabilities, highlighting the importance of Universal Design, which emphasizes opportunities to enhance accessibility of products for users with a broad range of abilities. The ERS [38] provides a widely used and well-validated measure of the ability for an environment to protect the rights, safety, and quality of life of people living with autism. It evaluates residential environments based on five subscales: structure, Communication, Socialization, Developmental Assessment and Planning, and Behavioral Management (see Table 1). The Functional Autonomy Measurement System (SMAF) [6] measures the level of independence of people living with autism and is stable and reliable, independent of the training level of the evaluator. This is a four-level, 29-item measure that rates a person's autonomy based on four subscales: activities of daily living, mobility, communication, mental functions and instrumental activities of daily living. For each of these subscales, the evaluator must rate the available resources that can assist the subject's disability in this area.

3 Designing for Independence

The CARE system has been designed as a supportive, low-cost, responsive environment that has the potential to empower independent living for adults with autism. Guided by the case studies (see section 4) and feedback from a transdisciplinary team of expert reviewers, from the fields of graphic design, architecture, human computer interaction, assistive technology, and autism support services, this paper reports on the development of a heuristic framework for the design of these assistive environments and interaction scenarios within them. CARE differentiates itself from previously designed smart homes by focusing specifically on providing a low-cost personally tailored option for increasing autonomous living, using off the shelf technologies. CARE takes the approach of providing a persuasive technology [12] that allows the person living with autism to either work independently or with a caregiver in order to personalize an environment tailored specifically to their needs. The ERS guided the approach to the development of scenarios and CARE/ERS heuristics and was used as a reference to solicit end-user feedback on the degree to which a given scenario or environment has the potential to accommodate the individual needs of persons living with autism.

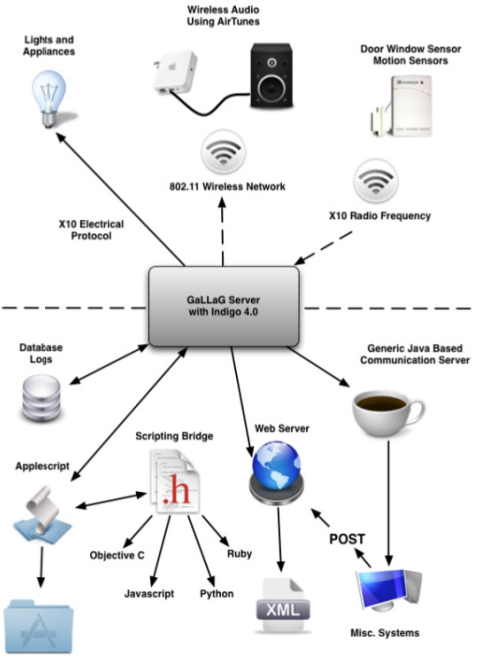


Fig. 1. The CARE system provides advanced e-health multimedia services in which physical in-home activities and interventions (top) are connected with interactive support, web-based services, distributed awareness, and remote intervention (bottom)

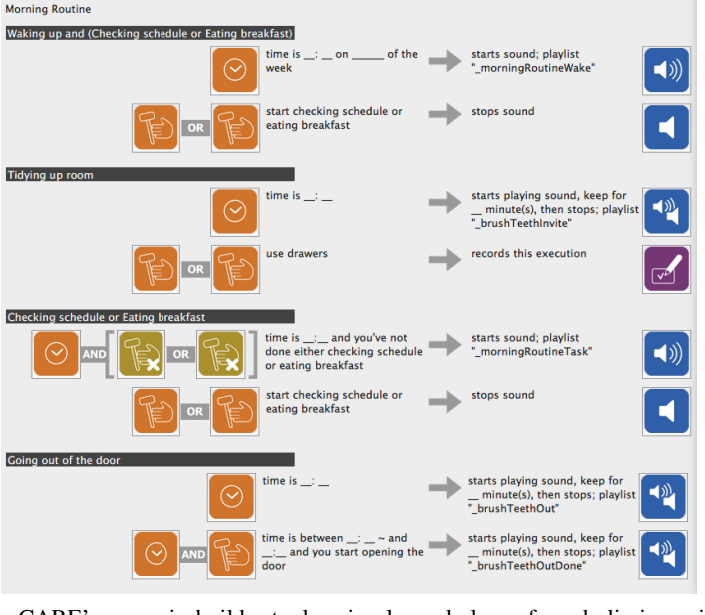


Fig. 2. Within CARE's scenario-builder tool a visual vocabulary of symbolic icons is used to co-develop, program, and customize personally tailored supportive scenarios

3.1 System Architecture

CARE's modular interaction architecture includes sensing, logging, and a system server (Figure 1). The system's architecture extends prior collaborations on the *Worker Interactive Networking* system employing a home sensing environment to assist adult children in caring for their elderly, live-in parents [25]. CARE advances a rich set of technologies for e-health, that can readily provide real-time multi-modal sensing and personally tailored supportive multimedia services and interventions (see section 4). Such sensing and actuation are provided through an extensible array of devices, including, but not limited to, X10 home automation components (door switches and motion sensors), RFID tags [32], wireless speakers, lighting actuators and sensors commonly found in mobile devices such as 3-D accelerometers, and GPS, all of which can be deployed throughout the home by end-users. The physical activities sensed by these devices transmit data to the Perspective Automation Indigo 4.0 server where it is logged into the database. The server facilitates the programing of interactive feedback, and communication (e.g., activity and time based logic, and audio/visual interactions), which when grounded in the context of user needs facilitates interventions. Likewise, feedback and communication are facilitated through the Indigo 4.0 server.

The Indigo 4.0 server is a robust home control server that communicates with INSTEON and X10 compatible hardware, such as televisions, light switches, etc., using their existing wiring arrangement. Indigo 4.0 uses a small fraction of the computer's CPU cycles so it does not significantly disrupt the availability or performance of an existing home computer, rather it augments it by extending its functionality beyond the traditional screen interface, to realize a low-cost off-the-shelf smart home environment. Collectively, these features of the CARE architecture facilitate end-user control of advanced e-health multimedia services that enable authorized co-located and distributed users (e.g., individuals, family members, caregivers, and medical professionals) access and control of scenarios and processes from remote locations, via webservers, and smartphones [25].

3.2 Scenario Design and User Experience

The process of configuring the CARE system accommodates the needs of people living with autism at each stage of the user experience, from scenario design and implementation, to the use of actualizing scenarios for activities of daily living, to the customized refinement of these scenarios (Figure 2). The process was designed for visual engagement, supporting people living with autism's tendency for positive responses to visual interaction [4]. The process allows users to drag and combine icons to create and program scenarios. Warm colors (yellow and orange) symbolize activity based triggers ("when the time is…" or "when I open this door…") while cool colors (blue and purple) symbolize resulting actions ("start playing music" or "record my action." This information is organized into two columns, which are separated by an arrow (this visual tool helps to indicate the relationship between the triggers and the resulting actions): triggers are on the left and resulting actions are on the right. To

evaluate the clarity of the symbolic programing system, prior to introducing it to adults living with autism and their caregivers, an expert reviewer, a senior communications director at a nationally ranked research university, was engaged to assess the clarity of the message of both the individual symbols and the "scenario-builder" as a whole. He confirmed the value of the color-coordination strategy and provided several suggestions. For example, he recommended that the "sound" icon be a different color than "no sound" to maintain consistency with the presentation of the "action" and "no action" icons which already employed two different colors. He also recommended that the readability of the "work" and "no work" icons be improved by representing them using the widely familiar construction worker symbol. Noting that the simplicity of the scenario-builder's tool was enhanced by adherence to the grid system, he also recommended limiting the use of text.

4 Case Studies

Iterative scenario-based refinement, involving participatory design by end-users [8], was employed to qualitatively evaluate and advance the CARE system. These human computer interaction methodologies, involving sketching user experience through participatory and co-design, are particularly important in the design of assistive and smart home technologies, due to the pervasive and intimately personal nature of homes. All facets of this human-subjects research were reviewed and approved by the Internal Review Board at [omitted]. Accordingly, to maintain participant anonymity, the names used below are pseudonyms.

The four case studies, presented here, draw from a series of ~1.5 hour meetings aimed at assessing and advancing CARE scenarios that appropriately attend to the most common needs and goals of people living with autism. Specifically, these meetings advanced, for both researchers and participants, an understanding of: how CARE can help people living with autism better meet their needs and pursue their goals; the process by which people living with autism can engage in CARE scenario creation and customization; and how CARE can be further developed to accommodate these scenarios within the homes of people living with autism. The meetings were conducted in three phases: (1) introducing the technology and its current and potential capabilities; (2) discussing the needs of the individual living with autism and sorting these needs and characteristics according to the DSM IV [4] autism spectrum disorder diagnostic criteria; (3) and engaging participants in the active design of their own scenarios. The 3 phase approach was informed by a subset of the ERS subscales: Phase 1 focuses on the *developmental assessment* Items 1, 2, 3, and 4; Phase 2 on *behavioral management* item 1; and Phase 3 on *communication* items 1, 2, and 3. While these specific items were the most relevant elements of the ERS with respect to the CARE case studies and system development process, each of the other items were also taken into account whenever applicable.

In Phase 2, based on DSM IV [4] autism spectrum disorder diagnostic criteria along with participant feedback, participants' characteristics were sorted into six

categories: social deficits; communication deficits; restricted interests (including lack of age-appropriate interests); repetitive patterns of behavior (including vocal or physical tics); problems with timeliness; and problems maintaining hygiene.

4.1 Case Study: Aaron

Aaron, an 18-year-old living with autism, aims to learn how to better interact with others so that he can "fit in with his classmates." His parents helped him break this long-term goal into shorter-term goals, which include increasing his social language skills, reducing his tics and jerky movements, and increasing his interest in age-appropriate activities. Together with Aaron and both of his parents, a researcher collaborated to develop four different possible scenarios for CARE in order to focus on his social deficits, repetitive behaviors, restrictive behaviors, and timeliness.

To address his social deficits, guided by ERS *socialization* subscale item 1, a scenario was designed in which Aaron can spin a digital/physical tetherball [34] that will randomly select either a person to Skype (he has a few friends that are willing to sign-up to be involved in his scenario), a virtual communication activity (using Sims 3, which he is currently using), or a word game. Within this scenario, each completed activity earns Aaron points (token economies [28] and competition have been shown to be effective motivators for boys living with autism [13]) that he can redeem for time on his computer. The physicality of the digital/physical tetherball allows the process to be a tangible, visually engaging experience [34].

To address his "repetitive behaviors" or jerky arm movements, guided by ERS *behavioral management* subscale item 5, a scenario was developed in which Aaron can attach wearable technology to his clothing when he is at home. This can sound a prompt if it detects sudden and/or repetitive movements (e.g., arm waving). CARE can also introduce regular audio cues, complemented with a visual representation of this schedule, either on his smart phone or projected onto his wall, to help Aaron conduct his daily activities in a more timely fashion, addressing ERS *structure* subscale items 1, 2, and 3.

A third scenario included activities that address his goal to reduce his "restrictive behavior" and to foster engagement in more age-appropriate interests that could help him better relate to his peers. Guided by ERS *communication* subscale items 3 and 5, s*tructure* subscale item 3, and *socialization* subscale item 1, his CARE scenario will sound an audio cue at 8pm to prompt Aaron to choose from a list of age-appropriate movies or activities.

Aaron's fourth scenario was developed to provide supportive direction throughout his day, guided by ERS *structure* subscale items 1, 2, and 3, CARE will sound an audio cue at 6am to wake him up, which be can turn off by opening his bedroom door. It sounds an audio cue at 7am to remind him to shave, which Aaron can silence by opening his bathroom drawer. At 8am, it sounds an audio cue telling him that it is time to get out the door. In order for the sound to stop and for Aaron to earn points for having his morning routine completed, Aaron must leave the house with his backpack, keys, and wallet, each of which are individually sensed by RFID tags.

4.2 Case Study: Bart

Bart, an 18-year-old living with autism, is a senior in high school who looks forward to studying film at [omitted] after taking a year off to live at home to get his driver's license and become more independent. His lack of timeliness makes waking up on his own, getting out the door on time, and remembering his wallet and ID very difficult for him. A researcher worked with him and his mother to design a scenario to motivate him to become more independent, guided by ERS *structure* subscale item 1 and ERS *behavioral management* subscale items 3, 4, and 5. At 6am, an audio cue wakes him up and will only stop if he enters the kitchen, where he is expected to make breakfast for himself and take pills. An audio cue will give him ten and fifteen minute warnings before the final reminder at 8:15am to get out the door. Guided by ERS *behavioral management* subscale item 5, when his backpack is removed from hanging on the wall and the front door opens, CARE will play music from his "Happy Music Playlist" on iTunes.

Bart's mother, his primary caregiver, designed this scenario more independently than the other participants, with less input from the researchers and Bart, both because she already had quite specific goals in mind for the system and because Bart became frustrated with his mother during the interview and left the room (his mother mentioned that he would probably live at the house for another year and he was upset because he wanted to live at college, though he did not explain why he left until after the interview with his mom was over). Before he left the room, a researcher explained the system and its goals and he did understand them and was willing to participate, but even when he was at the interview, his mother was more involved than he was. Bart and his family demonstrated some of the complex social relationships between adults with autism, their families, and their ability and motivation to participate in scenario design. Even prior to leaving the room, Bart seemed to trust his mother's decisions, so she evolved as the primary scenario-designer.

4.3 Case Study: Karen

Karen, a 19-year-old living with autism, has trouble maintaining her hygiene. Her dentist and parents have told her that she needs to brush her teeth more thoroughly; she almost always remembers to brush her teeth every morning and night, but when she does, she doesn't brush above and below the brackets on her braces. Working with Karen's parents, a scenario was designed in accordance with ERS *structure* subscale item 3, ERS *developmental assessment* subscale items 2 and 3, and ERS *behavioral management* subscale items 3, 4, and 5. When a motion detector in her bathroom senses that she is in the bathroom between the hours of 7 and 7:30am for her morning toothbrushing session and 8 and 8:30pm for her nighttime toothbrushing session, CARE will sound pleasant music from her iTunes playlist for two minutes (the chosen duration of her toothbrushing), offering reminders every 20 seconds (including "brush your lower teeth below the brackets," "brush your upper teeth above the brackets," "brush your lower teeth above the brackets," "brush your upper teeth below the brackets," etc.). If the time between the start and end of her toothbrushing

session is greater than two minutes, she will be awarded points that she can redeem with her parents for an iTunes purchase. While this may sound onerous, this is the scenario that was co-developed to meet Karen's individual needs.

Karen's mother was also interested in developing scenarios that would allow her to monitor Karen's behavior at school, where Karen often becomes upset. A scenario was developed in which CARE's iPhone integration could start video chats between Karen and her mother and her mother could text Karen's iPhone with feedback or advice. CARE could keep track of how many texts are sent and when they are sent. This would allow Karen and her mother to gain insight about what times of day are the most difficult for Karen so that they could work together to find a solution, guided by ERS *development assessment* subscale item 1.

4.4 Case Study: Jacob

Jacob, an 18-year-old living with autism, has trouble with timeliness. He often stays on the computer for so long that he forgets to do his chores, which include doing his homework, taking his dog out, and taking his medication and supplements. Guided by ERS *structure* subscale item 3, ERS *communication* subscale items 3 and 5, ERS *developmental assessment* subscale item 1, and ERS *behavioral management* subscale item 5, if Jacob is on the computer for more than an hour, CARE will sound an audio cue reminding him to do his homework. Once it senses that he is at his desk, the audio cue will play classical music while he works. This music will stop once it senses that the back door has opened and that Jacob is taking his dog out. When the door opens again, CARE reminds Jacob to take his medication. Guided by ERS *behavioral management* subscale item 5, once the pill bottle opens and closes, an activity detected by a small magnetic switch, CARE will provide a reinforcing audio feedback event, saying "Good Job!"

Jacob lives with his parents and has a caretaker that helps him with his social, communication, and timeliness skills during the week. Jacob hopes that CARE can help him become more independent as he begins to think about his future plans, though he is not sure if, when, or where he will move. His mother hopes that CARE can allow her and Jacob's caretaker more time to work on communication and social skills with Jacob, rather than wasting time and energy helping Jacob stick to his daily routine.

5 Assistive Smart Home Heuristics

Guided by each of the 32 ERS items (see Table 1), supported by heuristics introduced in related efforts [36, 39], and described in terms of the CARE system and scenarios, we have developed the CARE/ERS heuristic framework for smart home interaction design, implementation, and evaluation, for people living with autism. Of the various ways to assess the quality and usability of products for people living with autism, including Tsui's "Process for Developing Specialized Heuristics" [36], SMAF [6],

and ERS [38], ERS was considered most appropriate for the design and evaluation of personally tailored smart homes, because of its treatment-based holistic nature. This allowed us to not only evaluate the quality of CARE as a product or system, but also in terms of quality of life benefits to people living with autism. While research suggests that treatments based on the ERS result in improved quality of life and enhanced adaptation to new residential environments for adults living with autism [38], ERS does not accommodate assistive technology. This is in part because it is geared toward a group environment (2 or more people living with autism per residence) whereas CARE is centered on the individual. To address these shortcomings we developed the CARE/ERS heuristics that are applicable to technological supported assistive e-health interventions.

Table 1. Summary of CARE/ERS Heuristics

CARE Items	Corresponding ERS Item(s)
Communication Subscale	
1	1 Caregiver's language adjusted
2	2 Communication supplemented by visual systems
3	3 [Communication] training incorporated into daily routine
4	4 Information is available about communication skills
5	5 Directions communicated clearly to subject
Structure Subscale	
1	1 Physical organization of home facilitates independence
2	2 Visual systems used to teach new skills and maintain independence
3	3 Daily schedule for client visible 4 Daily schedule for home visible 5 Daily schedule is full and provides opportunities for leisure pursuits
Socialization Subscale	
1	1 Socialization training is incorporated into the daily routine 2 There are clearly stated goals for social, leisure, and affective skills 5 Social skills training involves skills used in interactions with others 6 Social skills are taught in meaningful contexts 3 Appropriate leisure activities are planned based on individual interests 4 Independent skills are developed for free time use
Developmental Assessment Subscale	
1	1 Caregivers are aware of [user's] cognitive level 2 Caregivers are aware of [user's] social level 4 Strengths and weaknesses incorporated into training activities 6 Functional needs are incorporated into training activities
2	3 Appropriate training activities selected based on assessment information 7 Training activities are rethought if the [user] is showing too little progress
3	5 Emerging skills are incorporated into training activities
4	8 Efforts are made to generalize skills
Behavioral Management Subscale	
1	1 Limits or rules are clear for the subject
2	2 Consistent strategies maintain behavioral limits
3	3 Problem behaviors are analyzed recognizing deficits and reasons behind the behavior 7 Written data is kept on behavior programs
4	4 [User] reinforced for positive behaviors 5 Behavior management strategies emphasize positive over punitive approaches
5	6 Less intrusive approaches are tried before more intrusive approaches

The CARE/ERS heuristic framework has undergone iterative review by expert reviewers (senior leaders in their fields involved in design, human computer interaction, assistive technology, and serving as directors of autism centers) in order to validate the framework and demonstrate CARE's potential as an assistive technology capable of improving living conditions and supporting aspirations. The following subsections will present the specific heuristics that culminated from this process, for each of the five ERS metrics in the context of CARE deployment scenarios. When CARE scenarios fulfill more than one of the ERS items, rather than repeating the scenario for each item, the multiple ERS items are aggregated to reflect the appropriate relationship.

5.1 Heuristics for Communication

Environments should have the ability to encourage effective communication between the user and the caretaker, as well as any other people the user may encounter. This CARE/ERS *communication* subscale has the following 5 items.

Item 1: *An assistive technology should be programmed to respond with verbal praise, and the wording should be specific rather than general.* For example, "Great job leaving for work on time!" is often a more effective and unambiguous reward than simply stating "Great job!" [29]

Item 2: *An assistive technology should be primarily visual*. When practical, a system should use visual aids to communicate to the user. This can be accomplished via projections of the user's schedule or calendar, supplemented with easily understood symbols and icons.

Item 3: *An assistive technology should incorporate communication into the daily routine.* Designers should make sure that scenarios incorporate communication training or opportunities for socialization.

Item 4: *Information about communication skills should be available.* For example, a library of communication tips and communication exercises specific to the user's needs may be incorporated.

Item 5: *All directions should be communicated clearly.* Each item on the schedule should be represented by an audio alert that the individual has chosen specifically for that task, and a visual representation of the schedule should be displayed.

5.2 Heuristics for Structure

The physical arrangement of the environment should suit the needs of the user. This CARE/ERS *structure* subscale has the following 3 items.

Item 1: *The physical organization of the system in the home should facilitate the individual's independence.* Even in situations where the system can do something for the individual, designers should consider whether it should be based on the individual's capabilities. For example, the system should not do an individual's laundry, but instead encourage him to learn to do it himself if it is within his capabilities.

Item 2: *Visual aids should be provided in order to assist in teaching new skills.* If an individual wants to cook a new meal, she should be provided with step-by-step instructions featuring pictures and words in addition to any auditory information.

Item 3: *The daily schedule should be customizable by the individual, encourage constructive activities, and be adaptive to accommodate unforeseen events.* When a user is creating her daily schedule, she should be encouraged to include events that promote physical health, such as a walk. If she misses an event, her schedule should adapt so that it reorganizes her task. The schedule of the user should be constantly visible and accessible, and capable of rearranging activities if the user misses an appointment or finishes one early.

5.3 Heuristics for Socialization

Environments should encourage the advancement of social activities and skills of the user. This CARE/ERS *socialization* subscale has 1 item.

Item 1: *Daily leisure activities should encourage socialization and skills development in meaningful contexts.* The system may recommend specific dates and times in an individual's schedule and request that they set up a social activity, such as a movie with a friend, at that time. If the user programs a social activity, such as a visit from a friend, into her schedule, the system should prioritize that, not rescheduling anything else in its place, and provide the individual with ample warning, through customized interventions (ie. an audio message stating, "Your friend is coming over in an hour! Are you dressed?").

5.4 Heuristics for Developmental Assessment

Environments should target users' acquisition of new skills, ensuring that goals set in place for them are both relevant and achievable. This CARE/ERS *developmental assessment* subscale has the following 4 items.

Item 1: *The user's strengths and weaknesses should be acknowledged and incorporated into the system.* Despite the additional time and effort, the person programming the system (whether that be the individual living with autism, a family member, or a caretaker) should be encouraged to pay close attention to the individual's interests and abilities in order to design a program that is tailored specifically to her needs.

Item 2: *If the individual is not making progress on a certain goal or becomes frustrated with the system, the system should make alternate options available.* If an individual is unable to complete her work on time, incremental rewards may be offered for productive steps rather than completion.

Item 3: *Emerging skills and new goals should be incorporated into the individual's training and scheduling.* Once an individual consistently gets to work on time, that goal should be replaced with a new goal (e.g., being fully prepared).

Item 4: *Efforts should be made to generalize skills.* For example, the system should recommend leisure activities that allow the individual to work on the social skills she learns in her online training.

5.5 Heuristics for Behavioral Management

Environments should be able to supervise and help regulate the user's conduct, as well as ensure that monitoring and instructive interventions are ethical and useful. While a system should allow people with autism as much control over the scenario design and use as possible, family members will often play important roles in these processes. Thus it is important to develop systems that not only address the needs of adults living with autism, but their caretakers as well. This CARE/ERS *behavior management* subscale has the following 5 items.

Item 1: *The conventions and directions for designing and using the system should be made clear to the individual with autism.* The individual should understand how to operate the system to the best of their ability.

Item 2: *The system should be reliable and free of "glitches."* Due to common sensory sensitivities, problems with the technology of the system can be highly intrusive and should be eliminated as effectively as possible.

Item 3: *The data that the system collects about the individual's task completion should be made available to the individual in order to analyze behavior.* If an individual is trying to increase the speed of her morning routine, she should be able to view a graph of her progress over the past month to see if it has improved.

Item 4: *The system should encourage positive reinforcement.* If the individual gets out the door on time, the system should offer an option to play happy music once she has done so.

Item 5: *Less intrusive approaches should be tried before more intrusive approaches.* The individual should first use a sound that will not upset them as an audio cue to wake up in the morning. Only if it does not work should the sound be changed to a more intrusive audio cue.

6 Conclusion

The CARE/ERS heuristic framework for smart home interaction design and the CARE system aim to increase independence and quality of life, contribute to the development of a broader set of options with increase capabilities to meet the needs of individuals, families, and caretakers. The CARE/ERS heuristics can guide the development of a broad array of smart home based assistive technologies and behavioral interventions. When implemented in accordance with the heuristics, CARE can offer support in a comfortable and unobtrusive manner that aims to increase quantity, quality, and affordability of residential options for people living with autism.

This investigation presents CARE as a low-cost, off-the-shelf ubiquitous computing interaction architecture that can be deployed as an empowering smart-home environment that provides supportive interactions that foster independence, learning, and

behavior modification. To contextualize the iterative design, a review of existing assistive smart home technologies and supportive technology designed for people living with autism has been presented. The iterative design of CARE scenarios have been firmly ground in the existing strengths of ERS. This has led to the creation and evaluation of a CARE/ERS heuristic framework for the design of ubiquitous assistive technologies for the residential environments of people living with autism. Collectively, these contributions present a participant centered design approach and heuristics that are broadly applicable to a new generation of multi-media residential e-health systems.

Acknowledgments. We thank Gregory Abowd, Kim Steele, Patrick Cheung, Sherry Ahrentzen, Jim Adams, the Arizona Health and Aging Lab at ASU, Southwest autism Research & Resource Center (SARRC), and most of all the case study participants for their contributions to this research.

References

1. Advancing Futures for Adults with Autism (2011), http://www.afaa-us.org/site/ (retrieved December 15, 2011)
2. Ahrentzen, S., Barger, T., Blackbourn, J., Bosworth, G., Gerhardt, P., Hannah, N., et al.: Opening Doors: A Discussion of Residential Options for Adults Living With Autism and Related Disorders. Autism Center (2009), http://www.autismcenter.org/documents/openingdoors_print_042610_001.pdf (retrieved December 15, 2011)
3. Ahrentzen, S., Steele, K.: Advancing Full Spectrum Housing: Designing for Adults with Autism Spectrum Disorders. Arizona Board of Regents, Tempe (2010)
4. American Psychiatric Association, Diagnostic and Statistical Manual of Mental Disorders IV, Text Revision, Washington, DC, p. 70 (July 2000)
5. Autism Spectrum Disorder. Center for Disease Control and Prevention (2010), http://www.cdc.gov/ncbddd/autism/index.html (retrieved on December 15, 2011)
6. Berberrt, R., Carrier, R., Bilodeau, A.: The Functional Autonomy Measurement System (SMAF): Description and validation of an instrument for the measurement of handicaps. Age and Ageing 17, 293–302 (1988)
7. Burke, M., Kraut, R., Williams, D.: Social use of computer-mediated communication by adults on the autism spectrum. In: ACM CSCW 2010: Conference on Computer-Supported Cooperative Work, pp. 425–434 (2010)
8. Buxton, B.: Sketching User Experiences: Getting the Design Right and the Right Design. Morgan Kaufman, CA (2007)
9. Carmien, S., Dawe, M., Fischer, G., Gorman, A., Kintsch, A., Sullivan Jr., J.: Socio-technical environments supporting people with cognitive disabilities using public transportation. ACM Transactions on Human Computer Interactions 12(2), 233 (2005)
10. Carroll, J.: Scenario-based design: envisioning work and technology in system development. John Wiley & Sons, Inc., New York (1995)
11. Cook, D.J., Youngblood, M., Heierman, E., Gopalratnam, K., Rao, S., Litvin, A., Khawaja, F.: MavHome: An agent-based smart home. In: Proceedings of Pervasive Computing 2003, pp. 521–524 (2003)

12. Fogg, B., Grudin, J., Nielsen, J., Card, S.: Persuasive Technology: Using Computers to Change What We Think and Do. Science & Technology Books (2002)
13. Geurts, H.M., Luman, M., Van Meel, C.S.: What's in a game: the effect of social motivation on interference control in boys with ADHD and autism spectrum disorders. Journal of Child Psychology and Psychiatry 49, 848–857 (2008)
14. Helfrich, D., Adrian, M.: Comfort, Clarity and Calm: Architecture for Autism. AUTISM Advocate 52(3), 1–6 (2008)
15. Heller, T., Caldwell, J.: Innovative project supports later-life planning for adults and their aging family caregivers. Maximizing Human Potential 8(3), 4–5 (2001)
16. Howlin, P., Goode, S.: Outcome in adult life for individuals with autism. In: Volkmar, F. (ed.) Autism and Developmental Disorders. Cambridge University Press (1998)
17. Indigo 4.0. Perceptive Automation (2011), http://www.perceptiveautomation.com/indigo/ (retrieved December 15, 2011)
18. Jabrink, K., Knapp, M.R.J.: The Economic Impact of Autism in Britain. Autism 5(1), 7–22 (2001)
19. Kientz, J., Patel, S., Jones, B., Price, E., Mynatt, E., Abowd, G.: The Georgia Tech aware home. In: CHI 2008 Extended Abstracts on Human Factors in Computing Systems (2008)
20. Krauss, M.W., Seltzer, M.M., Jacobson, H.T.: Adults with Autism living at home or in non-family settings: Positive and negative aspects of residential status. Journal of Intellectual Disability Research 49, 111–124 (2005); Anderson, R.E.: Social impacts of computing: Codes of professional ethics. Social Science Computing Review 10(2), 453–469 (1992)
21. Kraemer, B.R., Blacher, J.: Transition for young adults with severe mental retardation: school preparation, parent expectations, and family involvement. Mental Retardation 39, 423–435 (2001)
22. Kraemer, B., McIntyre, L., Blacher, J., James, S.: Quality of Life for Young Adults With Mental Retardation During Transition. Mental Retardation 41(4), 250–262 (2003)
23. Leroy, G., De Leo, G.: Mobile Communication and Data Gathering Software for Autistic Children and Their Caregivers. Paper presented at the Positive Design: Technology + Design + Management = Creating New Models of Possibility for All, Monterrey, Mexico (2008)
24. Madsen, M., El Kaliouby, R., Goodwin, M., Picard, R.: Technology for just-in-time in-situ learning of facial affect for persons diagnosed with an autism spectrum disorder. In: Proceedings of the 10th International ACM SIGACCESS Conference on Computers and Accessibility (2008)
25. Mahoney, D.: Linking Home Care and the Workplace Through Innovative Wireless Technology: The Worker Interactive Networking (WIN) Project. Home Health Care Management Practice 16, 417–428 (2004)
26. Mahoney, D., Mahoney, E., Liss, A.: AT EASE: Automated Technology for Elder Assessment, Safety, and Environmental monitoring. Gerontechnology 8(1), 11–25 (2009)
27. Matson, J.L., Benavidez, D.A., Compton, L.S., Paclawskyj, T., Baglio, C.: Behavioral treatment of autistic persons. Research in Developmental Disabilities 17, 433–465 (1996)
28. Matson, J.L., Boisjoli, J.A.: The token economy for children with intellectual disability and/or autism: a review. Research in Developmental Disabilities 30, 240–248 (2009)
29. McGee, G., Krantz, P., Mason, D., McClannahan, L.: A modified incidental-teaching procedure for autistic youth: acquisition and generalization of receptive object labels. Journal of Applied Behavior Analysis 16(3), 329–338 (1983)

30. MetLife Market Survey of Nursing Home, Assisted Living, Adult Day Services, and Home Care Costs (2010), http://www.metlife.com/assets/cao/mmi/publications/studies/2010/mmi-2010-market-survey-long-term-care-costs.pdf (retrieved on December 15, 2011)
31. Mitchell, P., Parsons, S., Leonard, A.: Using Virtual Environments for Teaching Social Understanding to 6 Adolescents with Autistic Spectrum Disorders. Journal of Autism and Developmental Disorders 37(3), 589–600 (2007)
32. O'Connor, M.: Linentracker Automates Management of Towels, Sheets. RFID Journal (2009), http://www.rfidjournal.com/article/view/533 (retrieved on December 15, 2011)
33. Olivier, P., Monk, A., Xu, G., Hoey, J.: Ambient kitchen: Designing situated services using a high fidelity prototyping environment, Corfu, Greece, 2 (2009)
34. Pendaphone (2011), http://www.pendaphonics.com (retrieved on December 15, 2011)
35. Teeters, A., El Kaliouby, R., Picard, R.W.: Self-Cam: Feedback From What Would be your Social Partner. In: ACM SIGGRAPH, 2006 Research Posters, Boston, MA, July 30-August 3, p. 138 (2006)
36. Tsui, K. M., Abu-Zahra, K., Casipe, R., M'Sadoques, J., Drury, J. L.: A Process for Developing Specialized Heuristics: Case Study in Assistive Robotics. Computer Science Tech Report 2009-11. University of Massachusetts Lowell, Lowell (2009)
37. Van Bourgondien, M.E., Mesibov, G.: Diagnosis and treatment of adolescents and adults with Autism. In: Dawson, G. (ed.) Autism, pp. 367–385. Guilford Press, New York (1989)
38. Van Bourgondien, G., Reichle, N., Campbell, D., Mesibov, G.: The environmental rating scale: A measure of the quality of the residential environment for adults with autism. Research in Developmental Disabilities 19(5), 381–394 (1998)
39. Vanderheiden, G., Vanderheiden, K.: Guidelines for the Design of Consumer Products to Increase Their Accessibility to Persons with Disabilities or Who Are Aging (1992), http://trace.wisc.edu/docs/consumer_product_guidelines/consumer.htm (retrieved on December 15, 2011)

Persuasive Sensing: A Novel In-Home Monitoring Technology to Assist Elderly Adult Diabetic Patients

Samir Chatterjee[1], Jongbok Byun[1], Akshay Pottathil[1], Miles N. Moore[2], Kaushik Dutta[3], and Harry (Qi) Xie[1]

[1] School of Information Systems & Technology, Claremont Graduate University, California, USA
[2] Advanced Warning Systems Inc., USA
[3] National University of Singapore, Singapore

Abstract. Diabetes mellitus is a common but serious chronic disease that kills thousands of patients worldwide each year. While there are several useful regimens that can be followed to manage the disease, elderly adult patients have particular difficulties in self-managing the disease. In this paper we present a novel approach to self-management – persuasive sensing – that uses environmental and body-wearable sensors that continuously detects activities and physiological parameters. Our system sends persuasive text messages and a weekly health newsletter aimed to alter the subject's behavior. We present the findings from an in-home monitoring implementation. The results obtained are quite encouraging. We discuss the challenges and lessons learned from such a field experiment and how we can improve upon the technology.

1 Introduction

Diabetes mellitus is the most common and serious chronic disease in the United States. There are nearly 26 million Americans with diabetes of which 10.9 million (or 26.9%) are aged 65 and older [1]. California in particular has the highest incidence of new diabetes cases and nearly 4 million people estimated to be suffering from the disease [2]. The costs of caring for this disease are astronomical and are estimated to exceed more than $24 billion in California and $174 billion nationally [1, 2].

Diabetes remains a major health problem being responsible for up to 8% of national health care expenditure. Diabetes mellitus is a chronic disease characterized by a sustained elevated blood glucose level, caused by a reduction in the action of insulin secretion where related metabolic disturbances generate severe, acute and long-term complications that are responsible for premature death and disability [10]. The World Health Organization projects that diabetes deaths will increase by more than 50% in the next ten years without urgent action. Most notably, diabetes deaths are projected to increase by over 80% in low-middle income countries between 2006 and 2015 [18].

Despite the availability of effective treatment, diabetes remains poorly controlled. Fewer than 7% of diabetic patients meet treatment goals for lipids, blood pressure,

M. Bang and E.L. Ragnemalm (Eds.): PERSUASIVE 2012, LNCS 7284, pp. 31–42, 2012.

and glycosylated hemoglobin A1c [3, 4]. Elderly patients with diabetes have higher rates of mortality, congestive heart failure, myocardial infarction and stroke as compared to age-matched controls without the disease [5]. Moreover, despite evidence that the mortality rate is decreasing over time, the rate of complications is remaining the same [6]. As a result, the average number of lifetime complications per patient is increasing as patients are living longer. With the incidence of diabetes rapidly rising, this is a fatal combination for the economic wellbeing of our health system.

Poor adherence to recommended self-management guidelines is well-recognized as a significant barrier to effective glycemic control. Improved outcomes have been associated with better adherence to medications, blood sugar self-monitoring, diet and lifestyle changes, and appointment attendance [7 - 9]. Barriers include time constraints, knowledge deficits, denial, limited social support, inadequate resources, and low self-efficacy.

Many of the elderly adults have difficulty achieving tight control because of the high degree of cognitive resources needed to manage diabetes. A major challenge in chronic disease self-management, particularly in elderly Americans, is social isolation [15]. Elderly diabetic patients with poor social support have twice the mortality rate of those with adequate support [16]. Furthermore, though families and friends want to help their loved ones better manage their conditions they often do not know how to act or when to act. Studies consistently show that patients with empowered caregivers or peers have better outcomes [16].

Mobile phones are an ideal platform for supporting chronic diseases like diabetes because they are ubiquitous, low-cost, reliable, real-time, and versatile; and unlike most technologies, actually enjoy greater usage amongst racial/ethnic minorities. In self-management, mobile phones can help individuals remember to do various health-related activities and record them, and also help others in their personal wellness ecosystem to review ongoing health patterns and respond quickly to changes in health status [11, 12]. Using wireless sensor networks within the home can help to remotely monitor activity of daily living (ADL) which can then be used to send effective reminders and feedback [13, 19].

In this paper we describe the design, implementation and evaluation of such a mobile and wireless sensor system. In particular we describe a novel approach that we call “persuasive sensing”, in which elderly patients with diabetes receive customized text messages based on their sensor data to motivate them towards a healthier lifestyle. They also receive a customized health newsletter (weekly) that is aimed to inform and educate them on their various daily activities. The prototype implementation and its evaluation show tremendous positive results.

2 System Architecture and Prototype

Any at-home healthcare solution must detect and respond to the activities and/or characteristics of the older person. A network of sensors (worn, carried, or environmental) is an ideal technology platform for detecting and responding to health-relevant parameters such as movement, sleep, weight, physiological data and social activity

[14]. In designing our system, the following key principles were kept in mind throughout the process:

- This is a healthcare problem, not a technology problem. At the center is the patient, not the technology. That also means as the experiment progresses, we must adapt based on patient's feedback.
- The simpler the technology, the better. Patients must comprehend what is being sent as feedback.
- Wireless Sensor Networks (WNSs) for healthcare are mission-critical; reliability is of paramount importance.
- The daily feedback persuasive messages must be kept fresh and not boring so the patient is eager to receive them and learn how to change his/her behavior.
- It has to work in the home, not just in the lab.

A WSN device is a packaged data collecting or actuating component, which includes a sensor and/or actuator, a radio stack, an enclosure, an embedded processor, and a power delivery mechanism [14]. The sensor interacts with the environment and sends an appropriate signal (analog or digital) to the embedded processor (also called microcontroller unit). We used Iris Mote technology developed by Intel and UC Berkeley labs. The mote hardware platform consists of a microprocessor and radio chip (MPR). Sensors connect directly to the mote processor radio boards via various interfaces. This combination gives the mote the ability to sense, compute and communicate. The mote enables raw data collected by the sensors to be analyzed in various ways before sending it to an aggregator (in our case a laptop) that we placed within the home. The aggregator then uploads daily activity data to the cloud through secured channels via the Internet. The following different types of sensors were implemented in this project:

- A simple on/off switch that detects open/close of garage door (through which subject leaves homes).
- A simple on/off switch that detects the back porch door for open or close.
- A blood glucose monitor device that can connect easily to the laptop via USB and which can upload Blood Glucose values daily.
- An infra-red analog sensor that detects presence in the bedroom.
- A photo sensor connected to the TV to detect television viewing.
- A pressure pad sensor (from Colonial Medical) placed in the couch in the living room in front of TV.
- A simple on/off switch to detect opening and closing of medication cabinet.
- A simple on/off switch to detect open/close of cabinet containing insulin.
- A wireless weight machine (from Tanita Corporation) that sends value via Bluetooth.
- A commercial body-wearable sensor from BodyMedia Inc. which is an arm-band that the patient wears 24 hours. This multi-sensor senses number of steps walked, quality of sleep, and many other physiological parameters. Data is sent to the cloud via USB upload.

The subject was shown how to log into BodyMedia website where he could input diet/nutrition information. Our system would then fetch daily diet data and compute total calories consumed. We also provided the patient with bottled water and asked him to only drink that during the course of the experiment. This was a simple way for us to monitor water intake.

The overall architecture schematic is shown in Fig. 1, while actual photos taken from the implementation site are shown in Fig. 2. The on/off switches were not wireless, so we had to run wires from them to the microcontroller. We originally had plans to sense the kitchen microwave and refrigerator usage also. But as the long wires would disturb the subject's mobility and accessibility in the kitchen, we abandoned that idea.

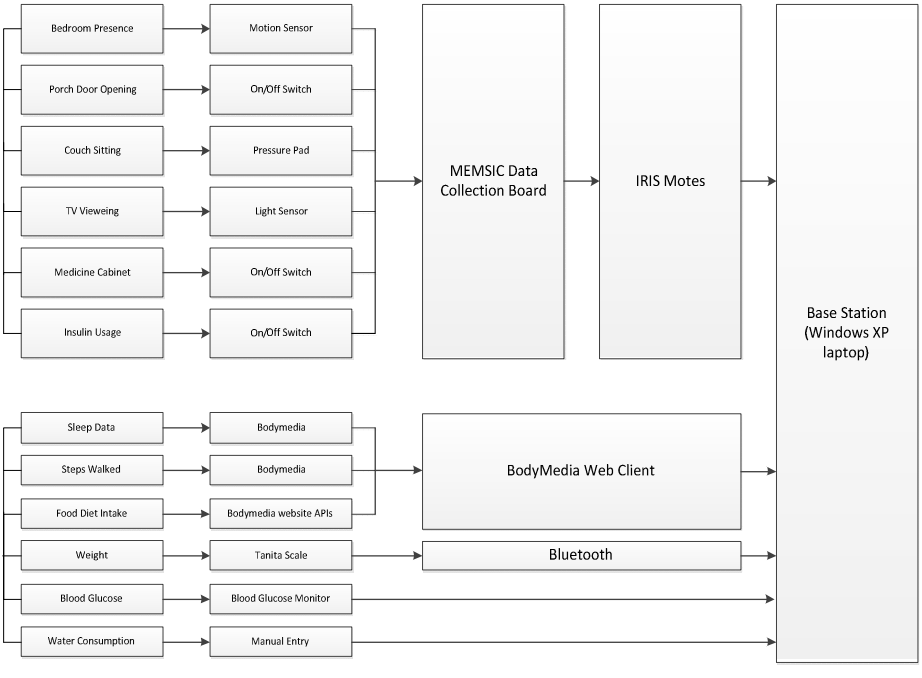


Fig. 1. Overall system architecture

3 Intervention Design and Persuasive Messaging Algorithms

Patients with type 2 diabetes can manage their chronic conditions by following certain recommended strategies. Prevention strategies for Type 2 diabetes include:

- Lose weight and keep Body Mass Index (BMI) under control
- 30 minutes or more of exercise or physical activity (brisk walking every day is fine)

- Develop a low calorie low fat diet. Nutrition guidelines include recommendations for a diet rich in whole grains, fruits and vegetables.
- Take necessary medications and measure their blood sugar regularly

Most elderly patients are challenged to adhere to these regimens due to lower cognitive abilities and lack of resources to maintain the lifestyle. Hence with technology, it is now possible to help these patients.

Our **intervention** has multiple components.

- System sends daily SMS text message on the cell phone using persuasive messages targeting behavior change.
- A tailored newsletter that summarizes healthy living parameters is presented to subject once a week and is jointly read by family member or one of our research team members.

Note our intervention (through the prototype persuasive sensing system) is aimed to engage patients in diabetes self-management through interactive SMS and newsletter approaches.

Fig. 2. Actual implementation photos from site (best viewed in color)

3.1 Primary Hypothesis

Hypothesis #1. Subject after the experimental intervention will show 5% reduction in blood glucose measurement (HbA1c measures 90-day average blood sugar).

Hypothesis #2. Subject will demonstrate more physical activity (measured by number of steps walked daily) after intervention and show trends that reach target goals set up before experiment.
Hypothesis #3. Subject will lower their daily calorie consumption and increase their intake of fruits and vegetables.

3.2 Secondary Hypothesis

Hypothesis #4. At end of intervention, subject will show improvement in BMI.
Hypothesis #5. At end of intervention subject will show weight loss.
Hypothesis #6. At end of intervention, subjects will demonstrate greater success in self-efficacy of diabetes management as measured by DES (Dissociative Experiences Scale).
Hypothesis #7. At follow up subject will show better quality of life.
It was important to ensure that daily text messages sent to the subject were fresh and relevant. Each day the subject received up to 3 text messages that were delivered to him over an LG smart phone. Below we show two examples of how messages were varied for physical activity and calorie consumption (Tables 1 and 2). The physical activity is measured by the number of steps obtained from the BodyMedia sensor.

Table 1. Messaging algorithm for physical activity

Case	Steps >= 8000	Steps < 8000
Mon	Great Job! Keep up the good work.	Don't give up on physical activity. Try walking a mile each day.
Tue	You have exceeded your goal. Congratulations.	Don't give up on physical activity. Have you taken the stairs?
Wed	You are doing very well. Keep it up!	Have you reached your goal of 8000 steps?
Thu	You are a super hero. You have exceeded your goal.	You fell short of your goal. Don't worry. Try to walk a mile after dinner.
Fri	You have exceeded your goal. Super job!	Never say never. You can do it.
Sat	Steps graph for past 5 days	
Sun	Great Job. Enjoy the Sunday with friends and family.	It is a beautiful day. Go out and do brisk walking for 30 mins.

To obtain diet/nutrition measures, the subject entered information into BodyMedia website. Using their APIs we were able to calculate total calorie consumed every day.

4 Subject Recruitment and Experiment Design

We obtained approval from our university Institutional Review Board (IRB). After IRB approved us to proceed, we distributed announcements to recruit subjects via

hospitals, diabetes clinics and through personal contacts. The basic eligibility criteria that we included in our recruitment efforts were:

- Subject must have Type 2 diabetes
- Age can be between 45-85
- Gender and race – no preference
- Have familiarity with cell phone and texting
- Have a broadband internet connection at home

Table 2. Messaging algorithm for food and nutrition

Case	cal <= 2500	cal > 2500
Mon	Your careful diet is going to help you reduce weight.	Try to eat reduced portions today!
Tue	Great job. You are watching your diet and will results soon.	Avoid takeaways and snack foods that are high in fat.
Wed	Enjoy green vegetables and salads and you are on your way to lose weight.	Try 10 baby carrots and a tablespoon of fat-free dressing for a 100-calorie snack.
Thu	Your diet calorie intake is under control. Congratulations!	Did you know that obesity is one of the leading causes of death in this country?
Fri	Very well done. If you can walk 5000 extra steps, then treat yourself a Star-bucks Frappuccino.	Have you tried low calorie drinks such as diet sprite?
Sat	Your diet trend is looking very good. Keep it up.	Add variety of colorful vegetables to your meal today.
Sun	Show calorie graph for past 6 days.	Choose low-fat dairy foods and lean meat.

We met with and interviewed prospective candidates who expressed interest. From the pool we selected our first subject. This subject is an 82 year old white male who is retired and lives in the Vista community near San Diego. He has type 2 diabetes, and a combination of other health problems. He agreed to the consent form and we started our project implementation.

We specifically designed a pre-post type of intervention.

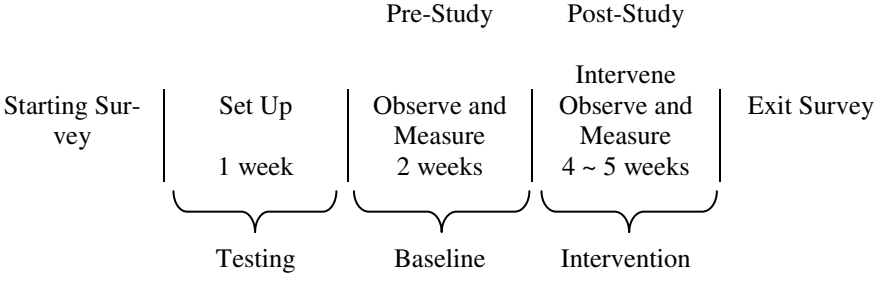


Fig. 3. Experiment design for intervention

We started our installation on October 18, 2011. Actual intervention through daily text messaging and weekly newsletter started from October 28, 2011. The experiment ended on November 25, 2011. All sensors and other equipment's were removed from the subjects' home after that. We thanked the subject and gave him a token honorarium for participation.

5 Results and Evaluation

As shown in Fig. 4A and Fig. 4B, the subject's blood glucose levels and weight is on decline trend. If subject can continue to reduce these physiological parameters, he will be in control of his chronic condition. Hypothesis #1 and #5 are supported. Both Fig. 4C and 4D demonstrate that subject has steadily increased physical activity and reduced his idle time throughout the intervention period. Hypothesis #2 is supported.

Fig. 4E shows that the subject's quality of sleep is improving in course of the intervention. Fig. 4F also shows that the sleep efficiency is getting better over time.

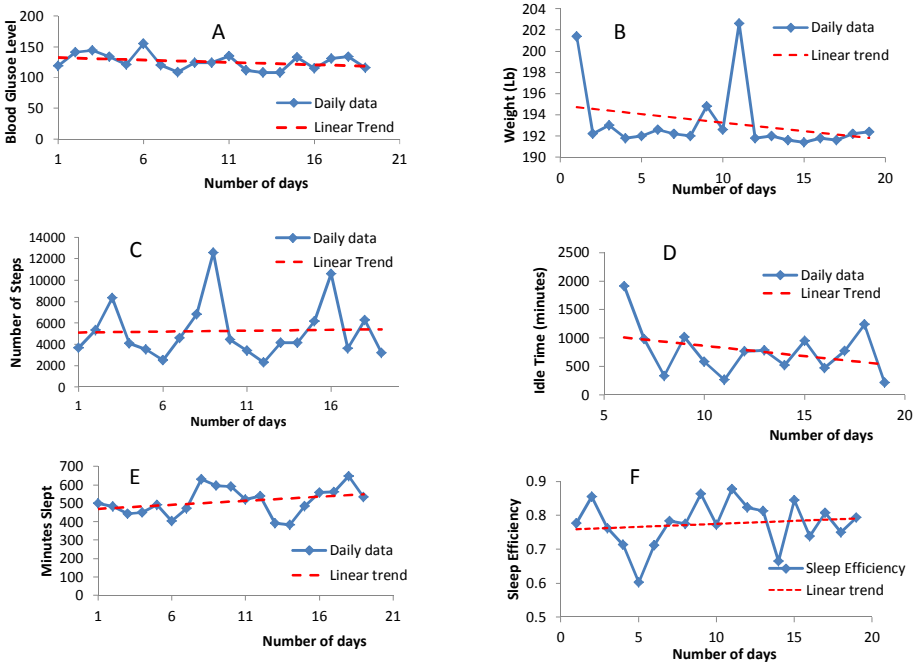


Fig. 4. Experiment results from the subject's home. (A) Blood glucose levels; (B) Weight trends; (C) Physical activity; (D) Idle time; (E) Minutes of good sleep; (F) Sleep efficiency

HbA1c is a lab test that shows the average amount of sugar in your blood over 3 months. It shows how well you are controlling your diabetes. An HbA1c of 6% or less is normal. Pre-experiment our subject's HbA1c was 12.9%. Post experiment HbA1c came to 6.6%. This is a significant improvement and validates hypothesis #1.

We also conducted a post-experiment exit survey in which we used the Diabetes DES scale to inquire about the subject's ability to manage the disease. Table 3. shows that subjects' ability to manage diabetes is improving (validates hypothesis #6) thereby improving his quality of life (hypothesis #7).

Table 3. Diabetes DES exit survey results

	Pre-Persuasive-sensing care	Post Persuasive-sensing care
I feel confident in my ability to manage my diabetes	agree	agree
I feel capable of handling my diabetes	agree	agree
I am able to do my own routine diabetes care.	agree	Strongly agree
I am able to meet the challenge of controlling my diabetes.	neutral	neutral

6 Analysis and Lessons Learned

The weekly health newsletter was customized for the subject. A snapshot of one page from the newsletter delivered on 11/24/2011 is shown below:

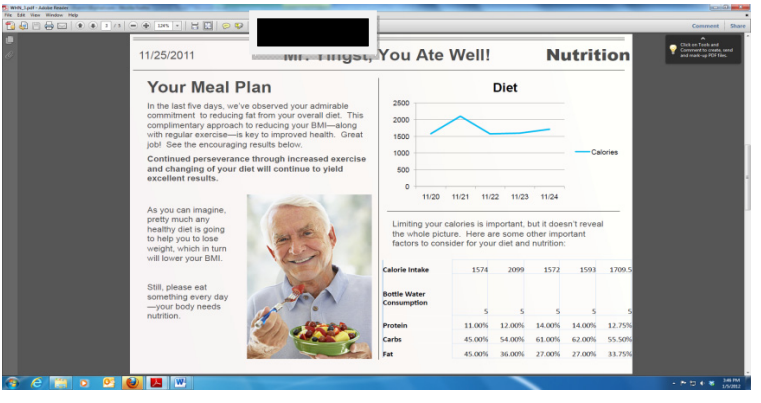


Fig. 5. Nutrition section of the newsletter

After each briefing of the newsletter with the subject, we received several useful feedbacks which helped us to customize the next version. During the implementation, we faced the following challenges:

- The Iris mote sensor system wasn't functioning sometimes and had reliability problems. A member of the team had to fix the errors several times before it became stable.

- Since on/off switches were not wireless, running wires from different parts of the home to the base station area was problematic. This can only be solved with wireless sensors and switches which we are deploying in the next implementation.
- The BodyMedia arm-band needs to be tight around the skin. However as our subject was elderly, it would not connect properly sometimes. Also at times it would give some skin rashes to the subject. A solution was to place the Body Media on his leg, and under his sock, which anchors the device against his leg, delivers a better measurement than placing the device on his arm.
- The base station being a Windows XP machine rebooted itself probably due to a security update and that killed many of the processes that we were using. A manual intervention was required in-between deployment to bring everything back to normal state.

The main feedback we received from each iteration of the newsletter was the following:

- Subject clearly understands the metrics associated with Physical Activity; but, does not have a clue as to what the "other key factors" mean. Suggestion: we should have a legend on this page to define the results, and better explain what they mean and how they correlate with Subject test objectives.
- Subject felt it did not capture the true amount of water consumed, and, he reiterated that he drinks voluminous amounts of water per day. Subject raised a question: is there something else you want me to take from this page in terms of this experiment?
- The Subject felt the Newsletter is a good tool for gathering relevant information important during physician visits/follow-ups; and, could be the basis for more meaningful dialog between Subjects and physicians.
- The 'persuasive' text messages are not bad; but, if a Subject like him sees the same ones time-and- time again, he gets numb and bored; so, the Subject recommends that messages should be specific – more personalized, catchy, and engaging in a narrative way to motivate the Subject to continue to modify behaviors and activities.
- The Subject observes that the food groups used in our study may be based on prepared food charts; and, does not account for naturally prepared food cooked fresh every day. With the exception of rice and pasta, the subject very rarely eats restaurant prepared food, or packaged meals.

As shown in Fig 6, we noticed that the subject's calorie intake increased steadily. For the demographic profile of the subject the goal calorie value should be around 2500 per day. So even though our hypothesis #3 was not supported, we felt that the subject was eating quite well and within the acceptable calorie levels.

A limitation of our project so far is that we are presenting results from one home and one subject. But we argue that there is a lot to be learned from this *case study*. The feedback we have received is invaluable. While the positive behavior change that

has resulted isn't statistically significant, we feel that our "persuasive sensing" idea has merit for further investigation. We have just completed a second home implementation (new subject is a 60 year old woman) whose HbA1c lowered from 8.9% to 8.5%. She has an iPhone and instead of a newsletter we used the smart phone capabilities and displayed the newsletter in a PDF format.

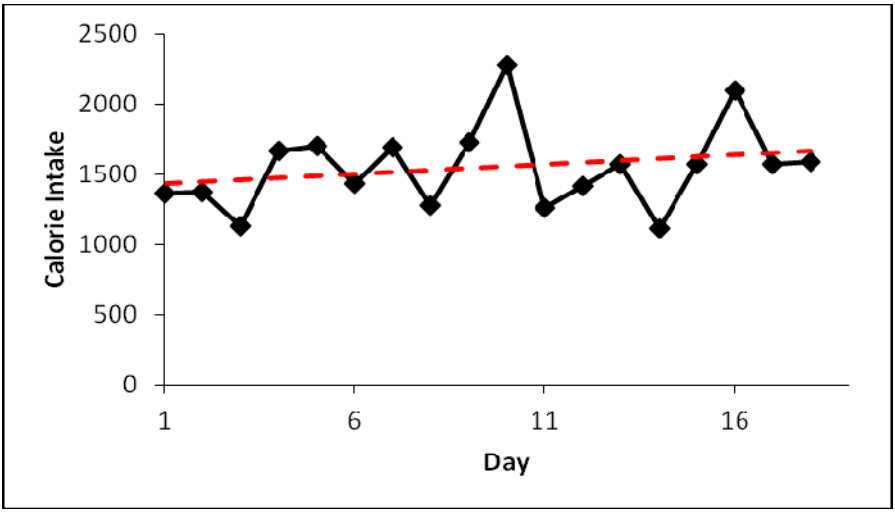


Fig. 6. Food/nutrition calorie trend

A concurrent activity of this project is to design artificial neural networks and machine learning algorithms that can mine the activity data and detect abnormal health conditions. Preliminary results were recently presented in [17] but we will report detailed results of that endeavor in a separate paper.

Acknowledgement. This project was partially supported by an EAGER grant from the National Science Foundation CNS Award #1048366.

References

1. Centers for Disease Control and Prevention. National diabetes fact sheet: national estimates and general information on diabetes and prediabetes in the United States. U.S. Department of Health and Human Services, Centers for Disease Control and Prevention, Atlanta, GA (2011)
2. California Diabetes Program, Diabetes Information Resource Center (DIRC). California Diabetes Fact Sheet. California Department of Public Health, University of California San Francisco (2011)
3. Saydah, S.H., Fradkin, J., Cowie, C.C.: Poor control of risk factors for vascular disease among adults with previously diagnosed diabetes. JAMA 291, 335–342 (2004)
4. Saaddine, J.B., Cadwell, B., Gregg, E.W., et al.: Improvements in diabetes processes of care and intermediate outcomes: United States, 1988-2002. Ann. Intern. Med. 144, 465–474 (2006)

5. Bethel, M.A., Sloan, F.A., Belsky, D., Feinglos, M.N.: Longitudinal incidence and prevalence of adverse outcomes of diabetes mellitus in elderly patients. Arch. Intern. Med. 167(9), 921–927 (2007)
6. Sloan, F.A., Bethel, M.A., Ruiz Jr., D., Shea, A.M., Feinglos, M.N.: The growing burden of diabetes mellitus in the US elderly population. Arch. Intern. Med. 168(2), 192–199 (2008); discussion 199
7. Schectman, J.M., Nadkarni, M.M., Voss, J.D.: Diabetes metabolic control and medication adherence in an indigent population. Diab. Care 25, 1015–1021 (2002)
8. Heisler, M., Faul, J.D., Hayward, R.A., Langa, K.M., Blaum, C., Weir, D.: Mechanisms for racial and ethnic disparities in glycemic control in middle-aged and older Americans in the health and retirement study. Arch. Intern. Med. 167, 1853–1860 (2007)
9. Blonde, L., Karter, A.J.: Current evidence regarding the value of self-monitored blood glucose testing. Am. J. Med. 118(9a), 20s–26s (2005)
10. Xiao, Y., Chen, H. (eds.): Mobile Telemedicine: A Computing and Networking Perspective, ch. 8. CRC Press, Taylor & Francis Group (2008)
11. Fogg, B.J., Adler, R. (eds.): Texting for Health: A simple powerful way to improve lives. Stanford Persuasive Lab. (2009)
12. Årsand, E., Tatara, N., Østengen, G., Hartvigsen, G.: Mobile Phone-Based Self-Management Tools for Type 2 Diabetes: The Few Touch Application. Journal of Diabetes Science and Technology 4(2) (March 2010)
13. Atallah, L., Lo, B., Ali, R., King, R., Yang, G.-Z.: Real-time Activity Classification Using Ambient and Wearable Sensors. IEEE Trans. Info Tech in Biomed. 13(6) (2009)
14. Dishongh, T.J., McGrath, M.: Wireless Sensor Networks for Healthcare Applications. Artech House (2010)
15. Liu, L.: Social connections, diabetes mellitus, and risk of mortality among white and African-American adults aged 70 and older: an eight-year follow-up study. Ann. Epidemiol. 21(1), 26–33 (2011)
16. Chesla, C.A.: Do family interventions improve health? Journal of Family Nursing 16, 355 (2010)
17. Chatterjee, S., Dutta, K., Xie, H., Byun, J., Pottathil, A., Moore, M.: Mining Behavior Profiles and Health Patterns from In-Home Monitoring of Sensor Data. In: Proc. of 21st Workshop on Information Technologies and Systems (WITS 2011), Shanghai, China, December 3-4 (2011)
18. World Health Organization. 10 facts about diabetes (2011), http://www.who.int/features/factfiles/diabetes/en/index.html
19. Virone, G., Alwan, M., Dalal, S., Kell, S.W., Turner, B., Stankovic, J.A., Felder, R.: Behavioral Patterns of Older Adults in Assisted Living. IEEE Transactions on Information Technology in Biomedicine 12(3) (May 2008)

Turning the Classic Snake Mobile Game into a Location–Based Exergame that Encourages Walking

Luca Chittaro and Riccardo Sioni

Human-Computer Interaction Lab
University of Udine
via delle Scienze 206
33100 Udine, Italy
http://hcilab.uniud.it

Abstract. Exergames (video games that combine exercise and play) could encourage physical activity by making it more enjoyable. Mobile devices are an interesting platform for exergames because they can support outdoors activities such as walking and running. Different mobile exergames have been proposed in the literature, and typically evaluated with informal interviews and ad–hoc questionnaires. The research we present in this paper had two main goals. First, we wanted to design a fun and easy-to-use mobile exergame to encourage walking. To this purpose, we propose a location–based version of the classic Snake mobile game, in which users can control the snake by walking. Second, we wanted to introduce important measures (such as users' attitude towards walking) in the evaluation of exergames, by adopting validated questionnaires employed in the medical literature. The results of the study presented in this paper shed light on how differences in users' lifestyle can be related to exergame enjoyment and to attitude change fostered by the exergame.

Keywords: Mobile games, Exergames, Location–based games, Attitude change.

1 Introduction

Regular physical activity can improve people's health and quality of life, lowering the risk of diseases like stroke and diabetes, improving cardiorespiratory and muscular fitness, decreasing levels of body fat and reducing symptoms of depression [1]. Moreover, for people who are inactive, even small increases in physical activity are associated with health benefits [1].

Video gaming is often blamed as a contributor to sedentary lifestyle, especially in children and teenagers. However, exergames, i.e., video games that combine play and exercise [2] such as Konami's Dance Dance Revolution, Nintendo's Wii Sports and Wii Fit, can make physical activity enjoyable for users, significantly increasing energy expenditure with respect to sedentary activities [3].

Mobile devices are a particularly interesting platform for designing novel exergames. Unlike desktop PCs and home consoles, mobile phones follow their users anywhere and are always available, allowing the creation of exergames that promote

M. Bang and E.L. Ragnemalm (Eds.): PERSUASIVE 2012, LNCS 7284, pp. 43–54, 2012.

outdoors physical activities such as walking and running. Different mobile exergames have been proposed in the literature (e.g., [4][5]). However, thorough and standard user evaluation practices are lacking. In particular, ad–hoc questionnaires are often used, which do not exploit the existing body of knowledge in measuring users' level of physical activity (e.g., [6]) or attitude towards physical activity (e.g., [7]).

The goal of our research was twofold. First, we wanted to design a mobile exergame that is fun and easy to understand. To this purpose, we decided to take the classic Snake mobile game (Fig. 1) and turn it into a location–based mobile exergame, which allows players to bring the snake into any real–world location and then control it by simply walking. The game (called LocoSnake, a blend of the words localization and snake) aims at engaging players in moderate outdoors physical activity. Seen as a persuasive technology, the game aims at encouraging walking and having a positive effect on the attitude of players towards walking.

Fig. 1. The classic Snake mobile game

Second, we wanted to set the evaluation of the exergame on a solid basis. To this purpose, we looked at the medical literature for possible standardized and widely adopted measures that could be employed to study mobile exergames and their effects in a reliable way. The study we present also focuses on possible differences in users' lifestyle concerning physical activity that might be related to game enjoyment and to the effects on attitude towards walking that the exergame might produce.

The paper is structured as follows. In section 2, we review the mobile exergame area and the evaluation techniques and instruments it has employed. Then, we describe LocoSnake (Section 3) and the method (Section 4) we employed in our evaluation. Section 5 and 6 respectively illustrate and discuss the findings of the study, while Section 7 illustrates conclusions and future work.

2 Related Work

Some of the existing mobile exergames run on dedicated devices. For example, Nintendo's Pokéwalker [8] is a small handheld device that includes a pedometer, which allows users to load a Pokémon character, put it in a pocket and take it with them for a walk, obtaining experience points and rare items.

Other mobile exergames rely instead on common smartphones, exploiting their internal sensors, such as accelerometers and GPS to build a stronger link between the game world and the real world. For example, GoldWalker [9] is an iPhone game

inspired by the California Gold Rush. By walking in the real world, players travel in the game world among cities, banks and mining camps, recruiting workers, buying tools and storing gold. Similarly, World of Workout [4] detects user's steps and wraps exergaming into a more conventional role–playing game. Players complete quests by walking for a given number of steps. In Hidden Park [10], players have to find magical creatures hidden in various locations in the real world by walking and solving puzzles. In Seek 'n Spell [11], a game for iPhone and Android phones, players walk in the real world to collect virtual letters located on the on–screen map. Points are earned as words are spelled. In PiNiZoRo [12], players have to find enemies by walking in the real world, playing various puzzle games to defeat them.

Another category of mobile exergames employs supplementary sensors to exploit additional user data, such as external accelerometers and heart rate (HR) sensors. Nike+ [13] for the iPhone, although not strictly a game, records the number of user's steps through a sensor inside Nike shoes, allowing the user to compare her activity with other users' and friends' data. Kazakos et al. [14] proposed some mobile exergames which employ an accelerometer attached to the user's waist to measure physical activity: for example, in NEAT-o-Race, when the user moves in the physical world, her avatar moves on a virtual racetrack, accumulating "activity points" and competing for them with friends. UbiFit Garden [15] relies on a wearable fitness device to detect and classify users' movements. It allows users to set weekly physical activity goals and, as they perform physical activities, it makes flowers bloom in a garden on the screen background. Butterflies appear when a goal is reached. Health defender [5], a Space Invaders–style game, reads users' HR and, at random intervals during the playing session, it requires users to reach the target HR shown on screen to gain bonus items. In Monster&Gold [16], HR is recorded while the user jogs outdoors, and game events are dynamically generated to keep the user in her optimal exercise range.

The mobile exergames surveyed above were evaluated in different ways. For some of them (generally commercial products), no evaluations are reported. Some of the others were informally evaluated through interviews after a brief gaming session or with more structured tasks and pilot trials which in a few cases lasted for days or weeks. For example, Health Defender [5] was informally evaluated by collecting users' opinions and remarks about the game, while PiNiZoRo [12] was evaluated by carrying out a pilot study with 4 adults who have children in the game target age, interviewing them after making them play with a pre–defined map, and then creating a custom one.

World of Workout [4] and Monster&Gold [16] were instead studied in more detail. Ten users played two missions in World of Workout. HR measurements were recorded before and after a gaming session, to check if the game provided a sufficient degree of physical activity. A 10–item Likert scale was then employed to assess participants' enjoyment and motivation. Monster&Gold was first evaluated on 14 users who played the game for up to 15 minutes with a questionnaire to assess game understandability and its motivational effects. A second evaluation was then carried out on an improved version of Monster&Gold using a more extensive questionnaire, focusing in more detail on the motivational effects of the game. NEAT-o-Race [14]

was evaluated with four different two–day sessions, in which 8 users played the game, each time adding new options. Finally, UbiFit Garden [15] was evaluated through a 3–week field trial, during which users were encouraged to wear the fitness device, carry the phone with them and confirm a daily activity list. Participants were interviewed at the beginning as well as during the field trial about their experience in the study.

In general, the reported evaluations showed positive results about exergames' effects on both motivation and amount of users' physical activity.

3 The Proposed Exergame

3.1 Game Design

In the classic Snake (Fig. 1), a virtual snake moves at constant speed inside a closed virtual playing field. Players have to direct the snake up, down, left or right to prevent its head from colliding with walls or with its tail. To get points, the snake has to eat on–screen items. This causes its tail to grow longer, requiring the user to be more careful in planning her movements.

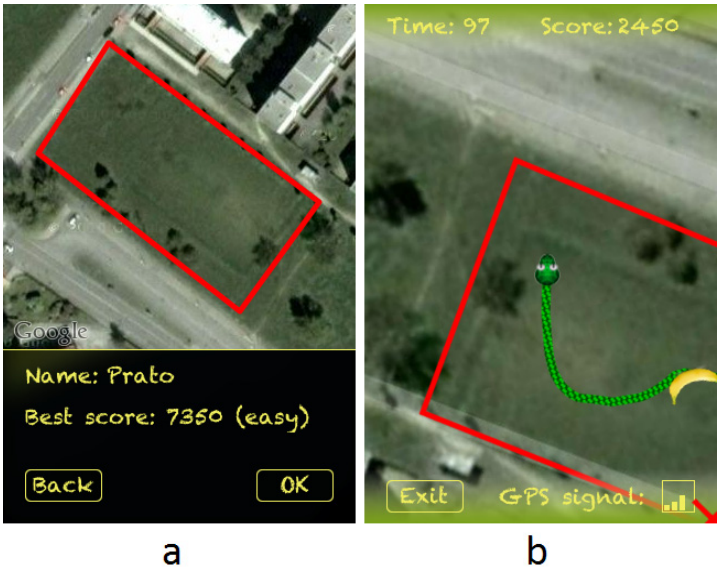


Fig. 2. Locosnake screenshots: (a) choosing the playing field; (b) playing (the red arrow at the bottom right angle points to an off–screen fruit)

In LocoSnake, the player embodies the snake and walks in the physical world to control it. Using GPS localization, the user can set the four–sided playing field (highlighted by red lines) on the visualized satellite map (Fig. 2a). Player's location in the real world determines the position of the snake head on the playing field (Fig. 2b).

To get points, the snake has to reach and eat virtual fruits located on the playing field. Unlike classic Snake, in LocoSnake the borders of the playing field can be hit by the snake without losing the game: if the GPS locates the player outside the playing field, the snake remains in the closest position inside the playing field. The player must instead be careful in avoiding collisions between the head and the body of the snake.

3.2 Game Setup

LocoSnake allows users to create and name their own playing fields (Fig. 2a). By creating their own playing field, users can play the game in any safe place they like, and by changing the playing field size as well as the difficulty level, users can change the intensity of the physical exercise, thus making the game more enthralling. Furthermore, the ability to play the game on a variety of different maps can allow enjoyment to last longer.

Users can choose among three difficulty levels, which change the time a fruit remains on the field: (i) Easy: a fruit disappears only when the snake eats it; (ii) Medium: a fruit, if not eaten, changes its position after a period of time proportional to the size of the playing field; (iii) Hard: analogous to medium, but fruit persistence time is halved. We related fruit persistence to the size of playing field in the medium and hard levels because users can create a playing field of arbitrary size: simply defining absolute time limits for eating a fruit could thus result in the game to be too hard with larger maps and too easy with smaller ones.

The map moves on the screen in such a way that the snake head is always at the center of the screen and map orientation is kept forward–up, so users can freely move around the field without the need to rotate the phone.

3.3 The Gaming Session

A level of the game is set to last 5 minutes. When playing, the time left and the current score are shown at the top of the screen. At the bottom, an “Exit” button allows the player to quit the game, and a small icon gives an indication of the accuracy of GPS data. Before or during a gaming session, the player can thus be warned about possible poor location data accuracy which is an issue for any location–based application. For this reason, before the user study, we tested the game for tens of hours in park lawns, and the level of GPS accuracy was always sufficient to successfully play the game.

There are always an apple and a banana simultaneously available inside the LocoSnake playing field. When the snake head comes close to a fruit, the phone vibrates, the snake shows its tongue and a biting sound is played. The fruit then disappears, the current score is updated and a new fruit appears in a random position inside the playing field.

When the game is over, it plays a tune and warns the user that level time has expired or the snake has bitten its tail. After that, a “Game Over” screen appears, showing the score and the distance in meters covered by the user.

4 User Study

The study we carried out focuses on three main aspects. First, we wanted to assess if LocoSnake is an enjoyable mobile game for users. Second, we wanted to evaluate the persuasive aspects of the exergame in terms of possible effects on attitude towards moderate physical activity, in particular walking. Third, we wanted to explore possible relations among users' lifestyle (in terms of physical activity), exergame enjoyment and possible effects on attitude towards walking. In the following, we describe the study in detail.

4.1 Design, Participants and Materials

Our study follows a within–subject design with walking activity (walking while playing LocoSnake and walking without playing LocoSnake) as the independent variable. For conciseness, we will refer to the first condition (walking while playing LocoSnake) as LS and to the other condition as NLS.

The evaluation involved a sample of 15 users (14 M, 1 F) recruited among computer science and mathematics graduate and undergraduate students at our university. Age ranged from 22 to 39 (M = 28.4, SD = 5.36). Video game usage in the sample was as follows: 2 participants never play video games, 5 participants play them a few times a month and 8 participants play at least a few times a week; 7 participants do not play exergames for PC or consoles, 5 participants play them about once a month, and 3 participants play them at least a few times a week; 14 participants do not play mobile exergames. Familiarity with GPS was as follows: 6 participants do not use GPS devices, 4 participants use them a few times a month, 5 participants use them a few times a week.

LocoSnake was run on an Apple iPhone 3GS equipped with a 3.5'', 320 x 480 pixels display. When playing LocoSnake, participants held the device in portrait orientation.

4.2 Measures

Game Enjoyment

To evaluate the enjoyment provided by LocoSnake, we employed 11 of the 13 items of the Player Enjoyment Scale (PES) [17], which derives from the EGameFlow scale [18] and is based on the concept of flow [19]. PES includes 13 bipolar statements, both positive and negative, with scores from 1 (strongly disagree) to 7 (strongly agree). We removed the 7th and 12th items ("The difficulty of challenges in the game increased as my skills improved" and "I experienced an altered sense of time while playing the game") because no increase of difficulty was planned (participants played a single level of the game) and the gaming session was short (5 minutes). We had to slightly change the 8th item ("The game provided new challenges with an appropriate pacing") into "The game proceeded with an appropriated pacing", because no new challenges were introduced during the single level played. This modified version of the scale (for brevity, mPES) returns a total enjoyment score in the 11–77 range,

obtained by reversing the negative item scores and summing them to positive item scores.

We also asked players four open questions about the exergame: what they liked and did not like about LocoSnake, if the game was difficult to use and if (and how) they would like to see it improved.

Perceived Exertion and Walking Speed

To measure the level of physical exertion perceived during the two experimental conditions, we employed Borg's Rating of Perceived Exertion (RPE) [20], a well–known and widely used instrument in the medical domain [21]. In this scale, users choose a number between 6 and 20 to describe their perceived exertion, being 6 "no exertion" and 20 "maximum exertion possible". We also computed the mean walking speed of participants in the two conditions. In the LS condition, the game recorded the distance covered by the participant using GPS data. In the NLS condition, we asked participants to keep the phone with them (without interacting with it) so that we were able to use GPS data for distance recording.

While we believed the game could receive good ratings in terms of player enjoyment, it was more difficult to formulate a hypothesis about the effects of the exergame on participants' perceived exertion and mean walking speed. On one hand, if the exergame is enjoyable and engaging, it can greatly absorb users' attention and distract them from the walking activity: this could lead users to perceive a lower exertion level and possibly walk faster to achieve the game goals. On the other hand, the continuous attention required by the game as well as the need to multitask in order to control walking and achieve game goals at the same time might actually cause the walking activity to be perceived as more fatiguing than a condition in which one can focus on simply walking.

Physical Activity Enjoyment and Attitude Change

To compare user's attitude towards walking in general and walking with LocoSnake, we employed the Physical Activity Enjoyment Scale (PACES) [7], an instrument thoroughly validated in the sports and exercise literature (e.g., [22]). This scale includes 18 bipolar statements with scores from 1 (strongly disagree) to 7 (strongly agree). The PACES produces a total enjoyment score in the 18–126 range.

In our study, each participant filled a PACES at the beginning of the evaluation concerning her general perception of the walking activity, and a PACES after playing LocoSnake concerning her perception of the walking activity when playing the exergame. We hypothesized that, by adding a gaming element to the physical activity, the exergame could have an effect on user's attitude, improving users' enjoyment of walking. We thus expected the score for the second PACES to be higher than the score for the first PACES.

Level of Physical Activity in User's Lifestyle

To assess level of physical activity in user's lifestyle, we employed the standardized Global Physical Activity Questionnaire (GPAQ) [6]. The GPAQ contains 16 questions that collect information on physical activity participation in three settings (activity at work, travel to and from places and recreational activities), as well as

sedentary behavior. After defining the MET (metabolic equivalent) as the energy cost of sitting quietly, the GPAQ analysis guide [6] specifies how to calculate the weekly amount of physical activity of a person in terms of MET–minutes per week (METM), i.e., the weighted time spent on moderate– and vigorous–intensity activities. The METM value is an indicator of how much a person is physically active.

We hypothesize that people with sedentary lifestyles could respond better to an exergame such as LocoSnake rather than more active people, in terms of game enjoyment as well as positive effects on attitude towards physical activity. For active people who are already motivated to walk, playing a game while walking might indeed be perceived more as added cognitive load rather than a pleasant distraction.

4.3 Procedure

We met each user at our university and guided her/him to a nearby lawn. Users were told that the goal of the study was to evaluate a game for the iPhone. They filled a demographic questionnaire to record age, gender and current occupation, video game usage, desktop exergames usage, mobile exergames usage and GPS devices and software usage. Then, they answered the GPAQ questions and filled the first PACES.

Each user was asked to walk while playing LocoSnake (LS condition) and to take a free walk around the lawn (NLS condition). In both conditions, users were asked to walk normally. To avoid possible confounds, in the NLS condition users could not carry out other activities (e.g., listening to music or talking on the phone). In this way, any difference in the observed results could reasonably be ascribed to the use of LocoSnake. The order of the two conditions was varied, so that for 7 out of 15 users the LS condition preceded the NLS one.

Before the LS condition, the experimenter showed the user the game and briefly explained how it works, then let her play the game for about 1 min for familiarization purposes. Then, the LS condition started and the user played a 5–minutes session with LocoSnake at the “Easy” difficulty level. After the session, users filled the RPE scale, the second PACES and the mPES. Finally, they could answer the four open questions.

In the NLS condition, the experimenter asked users to simply keep the iPhone (which recorded GPS data) with them, without interacting with it, and walk a lap around the lawn. Users were stopped after 5 min of walking and filled the RPE scale.

Finally, users were debriefed and thanked for their participation.

5 Results

5.1 Means

The reliability of the mPES was confirmed by calculating Cronbach’s alpha (α=0.74). The mean mPES score obtained by LocoSnake was 60.13 (SD = 6.58).

Wilcoxon test revealed that the difference between the RPE scores for the LS condition (M=9.07, SD=2.05) and those for the NLS condition (M=7.80, SD=1.52)

was statistically significant ($W=59$, $p < 0.05$): users perceived a greater exertion when playing with LocoSnake than when walking without playing the game.

Average walking speed was slightly faster in the LS condition than NLS (1.62 m/s and 1.45 m/s), but a t–test revealed that the difference was not statistically significant.

The mean score for the second PACES ($M=99.53$, $SD=9.29$) was slightly lower than the score for the first PACES ($M=102.07$, $SD=12.52$) and the difference was not statistically significant. The difference between the PACES scores (ΔPACES) was negative for 8 participants, positive for 5, and zero for 2 participants.

5.2 Correlations

Spearman's test revealed a negative correlation between METM values and mPES scores ($\rho(15) = -0.66$, $p < 0.01$), a negative correlation between METM values and ΔPACES scores ($\rho(15) = -0.59$, $p < 0.05$), a positive correlation between mPES and ΔPACES scores ($\rho(15) = 0.56$, $p < 0.05$. Fig. 3 graphically summarizes the correlations: users with greater METM values (i.e., who have a more active lifestyle) tend to enjoy LocoSnake less than users who are more sedentary, and the difference in attitude towards walking with and without the exergame (ΔPACES) is lower with more active users. Finally, users who find the exergame more enjoyable (mPES) are associated with higher values for the obtained difference in attitude.

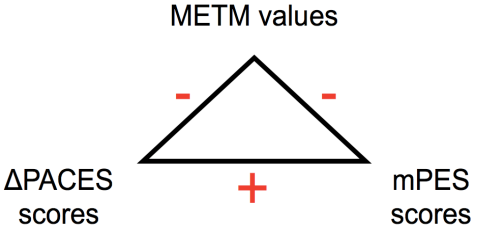


Fig. 3. Correlations among variables

5.3 Answers to Open Questions

Users gave the following answers to the question "What did you like about LocoSnake?": (i) The link between the game and the real world (6 users); (ii) Feedback provided by the game (4 users) such as real–time map rotation, vibration when eating a fruit, arrows pointing to off–screen fruits; (iii) the connection between physical activity and video gaming (3 users); (iv) The originality of the idea (3 users). Answers provided by single users reported liking for: similarity with the original Snake game; relaxing music; graphics; fun; challenge; ease of use.

In answering the question "What did you dislike about LocoSnake?", 9 users mentioned GPS–related problems, i.e. GPS lag and/or inaccuracy. Answers provided by single users reported disliking for: graphics, colors of the map, background music.

Five users replied to the question "Did you experience difficulties in using the game?": they reported difficulties related to GPS inaccuracy such as "the snake was moving jerkily" or "sometimes the snake had difficulties in following me".

Users gave the following answers to the question "What would you change to improve the LocoSnake game?": (i) improve position accuracy (7 users); (ii) add more types of fruits (2 users); (iii) show the score given by each fruit in a better way, e.g. by making it flash for a moment over the eaten fruit (2 users). Answers provided by single users suggested to: make the tail grow faster and longer, add different colors for the snake, add side effects to some fruits (e.g. slowing down the snake), add virtual obstacles on the playing field, add multiplayer, provide an alternative character and items (e.g., a car running over zombies).

6 Discussion

As indicated by mPES scores, users generally enjoyed the exergame. The fact that LocoSnake is based on a well–known mobile game helped user to immediately understand how to play: all users gave the highest rating to the 4th item in the mPES scale ("Overall, game goals were presented clearly"). Moreover, the answers to the last two items ("I temporarily forgot worries about everyday life while playing the game" and "I felt emotionally involved in the game") received a mean score of 5.47 (SD=1.46) and 5.27 (SD=1.58). Answers to the first open question (as seen in the previous section) provide some specific reasons for users' liking.

Users perceived a higher physical exertion in the LS condition with respect to the NLS condition. Some users reported "a greater exertion" because they "felt very engrossed in reaching the fruits", and perceiving that they were "walking fast" while playing LocoSnake (although in general average speed was not significantly different between the two conditions). We also observed that walking speed in the LS condition was higher when the fruit was far from the user, while it was lower when the user was closer to the fruit. This indicates that users modulated their speed while carefully navigating the physical environment as a strategy to accurately get the fruits in a short time. This could also have contributed to the perception of higher physical exertion.

A clear result of the open questions is that many users felt that the snake did not move as precisely as they expected, as confirmed by the mean scores obtained by the mPES items "I received immediate feedback on my actions in the game" and "I felt a sense of control over the game", respectively 3 (SD=1.13) and 4.73 (SD=1.39).

Unexpectedly, the difference between the scores of the second and the first PACES was positive only for a minority of participants. However, the correlations among METM, ΔPACES and mPES reported in the previous section seem to suggest that the mobile exergame can be more successful as well as more useful for people who have more sedentary lifestyles. Interestingly, two of the most physically active users provided the same explanation for their negative ΔPACES: they reported that walking is for them an activity in which "the mind can be left to wander" and the focus required by the exergame made thus the activity less pleasant.

7 Conclusions and Future Work

In this paper, we proposed a mobile exergame aimed at encouraging outdoor walking and we carried out an in–depth user evaluation, employing instruments proposed by international organizations (GPAQ) or whose reliability was thoroughly assessed in the literature (PACES and PES). The results suggest possible relationships among users' lifestyle, exergame enjoyment and the effects on the attitude towards walking that the exergame could bring. It would now be important to replicate the study with female participants (our sample was mostly male) as well as with other exergames, to test if the correlations can be generalized.

We are currently interested in introducing multiplayer features into LocoSnake, allowing multiple snakes to compete on the same playing field. Adding a social dimension to the game might indeed make the exergame more interesting and engaging also for physically active people. As discussed by Mueller et al. [23], having other human beings join a mediated exertion activity can positively contribute to the experience.

A limitation of the study is that it did not evaluate long–term effects of the exergame. Klasnja et al. [24] point out that longitudinal studies and randomized control trials of persuasive technologies for encouraging physical activity are difficult to carry out and not always feasible [24], but they also provide some recommendations to evaluate these technologies over time that we will consider for future studies. For example, a new version of LocoSnake could record the distance walked by the player during the various gaming sessions, as well as the time spent playing the game, keeping an automatic diary that could be useful to the user (to support a self–monitoring persuasive strategy as in UbiFit Garden [15]) as well as to the experimenter (to analyze changes in user's behavior over time).

References

1. US Department of Health and Human Services, http://healthypeople.gov/2020/topicsobjectives2020/overview.aspx?topicid=33
2. Bogost, I.: The rhetoric of exergaming. In: 42nd Design Automation Conference (2005), http://www.bogost.com/downloads/I.%20Boogst%20The%20Rhetoric%20of%20Exergaming.pdf
3. Graf, D.L., Pratt, L.V., Hester, C.N., Short, K.R.: Playing Active Video Games Increases Energy Expenditure in Children. Pediatrics 124(2), 534–540 (2009)
4. Doran, K., Pickford, S., Austin, C., Walker, T., Barnes, T.: World of Workout: Towards Pervasive, Intrinsically Motivated, Mobile Exergaming. In: Meaningful Play 2010 Conference, Michigan State University (2010)
5. Wylie, C.G., Coulton, P.: Mobile Exergaming. In: 2008 International Conference on Advances in Computer Entertainment Technology, pp. 338–341. ACM Press, New York (2008)
6. World Health Organization: Chronic diseases and health promotion — Global Physical Activity Surveillance, http://www.who.int/chp/steps/GPAQ/en/index.html

7. Kendzierski, D., DeCarlo, K.J.: Physical Activity Enjoyment Scale: Two Validation Studies. J. Sport Exerc. Psychol. 13(1), 50–64 (1991)
8. Pokéwalker, http://www.pokemongoldsilver.com/us/#/features/pokewalker/
9. GoldWalker, http://itunes.apple.com/us/app/goldwalker/id372683234?mt=8
10. The Hidden Park, http://www.thehiddenpark.com/
11. Seek 'n Spell, http://www.seeknspell.com/
12. Stanley, K.G., Livingston, I., Bandurka, A., Kapiszka, R., Mandryk, R.L.: PiNiZoRo: A GPS–Based Exercise Game for Families. In: 2010 International Academic Conference on the Future of Game Design and Technology, pp. 243–246. ACM Press, New York (2010)
13. Nike+, http://nikerunning.nike.com/nikeos/p/nikeplus/en_US/plus/#//dashboard/
14. Kazakos, K., Fujiki, Y., Puri, C., Bhuddaraju, P., Pavlidis, I., Levine, J.: NEAT-o-Games: Exertion Interface Interwoven in Daily Life. In: 26th ACM Conference on Human Factors in Computing Systems Workshop on Exertion Interfaces, pp. 1–3. ACM Press, New York (2008)
15. Consolvo, S., McDonald, D.W., Toscos, T., Chen, M.Y., Froehlich, J., Harrison, B., Klasnja, P., LaMarca, A., LeGrand, L., Libby, R., Smith, I., Landay, J.A.: Activity Sensing in the Wild: A Field Trial of UbiFit Garden. In: 26th ACM Conference on Human Factors in Computing Systems, pp. 1797–1806. ACM Press, New York (2008)
16. Buttussi, F., Chittaro, L.: Smarter Phones for Healthier Lifestyles: An Adaptive Fitness Game and Its Evaluation. IEEE Pervasive Computing 9(4), 51–57 (2010)
17. Chu Yew Lee, S.L., Been-Lirn Duh, H., Quek, F.: Investigating Narrative in Mobile Games for Seniors. In: 28th ACM Conference on Human Factors in Computing Systems, pp. 669–672. ACM Press, New York (2010)
18. Fu, F.L., Su, R.C., Yu, S.C.: EGameFlow: A Scale to Measure Learners' Enjoyment of E–Learning Games. Computers and Education 52(1), 101–112 (2009)
19. Csikszentmihalyi, M.: Flow: The Psychology of Optimal Experience. Harper Collins, New York (1991)
20. Borg, G.: Perceived Exertion as an Indicator of Somatic Stress. Scand. J. Rehabil. Med. 2(2), 92–98 (1970)
21. U.S. Centers for Disease Control and Prevention, http://www.cdc.gov/physicalactivity/everyone/measuring/exertion.html
22. Croker, P.R.E., Bouffard, M., Gessaroli, M.E.: Measuring Enjoyment in Youth Sport Settings: A Confirmatory Factor Analysis of Physical Activity Enjoyment Scale. J. Sport Exerc. Psychol. 17(2), 200–205 (1995)
23. Mueller, F., Vetere, F., Gibbs, M.R., Agamanolis, S., Sheridan, J.: Jogging Over a Distance: The Influence of Design in Parallel Exertion Games. In: 5th ACM SIGGRAPH Symposium on Video Games, pp. 63–68. ACM Press, New York (2010)
24. Klasnja, P., Consolvo, S., Pratt, W.: How to Evaluate Technologies for Health Behavior Change in HCI Research. In: 29th ACM Conference on Human Factors in Computing Systems, pp. 3063–3072. ACM Press, New York (2011)

Phone Row: A Smartphone Game Designed to Persuade People to Engage in Moderate-Intensity Physical Activity

Matthijs Jan Zwinderman, Azadeh Shirzad, Xinyu Ma, Prina Bajracharya, Hans Sandberg, and Maurits Clemens Kaptein

User System Interaction, Industrial Design Department, Eindhoven University of Technology, P.O. Box 513, Eindhoven 5600 MB, The Netherlands
{m.j.zwinderman,a.shirzad,x.ma,p.bajracharya, h.sandberg,m.c.kaptein}@tue.nl

Abstract. Few people reach the recommended levels of moderate-intensity physical activity (MIPA). This study examines whether persuasive technology, in the form of a smartphone game, can help people engage more in MIPA. A smartphone boat racing game was developed that requires users to make rowing movements and therefore engage in MIPA to play it successfully. With these rowing movements, users can control the movement of a virtual boat across a virtual track on an external screen. Users were fond of the concept of the game. However, a sub-optimal implementation resulted in users not wanting to replay the game and thus not developing a habit involving performance of MIPA. The implementation of the concept was inadequate for testing the hypothesis that a smartphone game can help people engage more in MIPA.

Keywords: moderate-intensity physical activity, exercise, smartphone game, persuasive technology, exergaming.

1 Introduction

The goal of the current study was to examine whether persuasive technology, in the form of a physically active smartphone game, can help people engage more in moderate-intensity physical activity (MIPA).

1.1 Moderate-Intensity Physical Activity

Moderate-intensity physical activity can be defined as an aerobic activity that can be performed while maintaining an uninterrupted conversation [1]. Examples are cycling, walking and active commuting [2].

Various studies have linked MIPA to health benefits. A meta-analysis of 22 studies, in which almost a million people participated, looked at the relationship between MIPA and all-cause mortality risk [2]. The meta-analysis found that 2.5 hours a week of MIPA, compared with no physical activity, is associated with a reduced mortality risk of 19%, while seven hours a week of MIPA reduces the mortality risk by 24% [2]. Another study found that even a low amount of MIPA is beneficial for a person's health. This study found that people that engage in MIPA for an average of 92

M. Bang and E.L. Ragnemalm (Eds.): PERSUASIVE 2012, LNCS 7284, pp. 55–66, 2012.

minutes per week had a mortality risk that was 14% lower and a three year longer life expectancy than people that engage in almost no MIPA [4].

All over the world, only a minority of people reach recommended weekly amounts of MIPA. In the Netherlands, 30 minutes of MIPA every day is recommended, but less than half of the Dutch population reaches this amount [5]. The World Health Organization [6] and the U.S. department of health and human services [7] recommend that healthy adults aged 18-65 years should engage in at least 150 minutes of MIPA per week. In the U.S., a third of adults aged 18-65 reach this recommendation [8]. In Taiwan, less than 15% reach the recommendation of at least three sessions per week of at least 30 minutes of MIPA [9].

1.2 Persuasive Technology and Physical Activity

The goal of the current study was to examine whether persuasive technology, in the form of a smartphone game, can help people engage more in moderate-intensity physical activity (MIPA).

Persuasive technology is "a class of technologies that are intentionally designed to change a person's attitude or behavior" [10]. Persuasive technology is used in many fields, such as cessation of smoking, reduction of food-consumption, increase of workouts and reduction of energy consumption in families [10] [11]. Strategies used to persuade people are, for instance, social proof ("your friends like it"), scarcity ("only two items remaining"), liking ("the attractiveness of the behavior"), reciprocity ("I'll scratch your back, if you scratch mine"), praise, authority, and commitment and consistency (i.e. developing habits) [12] [13].

Step Counters and Self-Reporting

Existing persuasive technologies designed to make people more physically active have generally focused on all-day physical activity, by using step counters or self-reporting (see e.g. [15] or [16]). Examples of available commercial products with these functions are Philips Directlife[1], Mytrak[2], Fitbit[3], Fitbug[4], Bodybugg[5], and Zamzee[6]. Besides the measuring device, these companies offer services such as tracking physical activity levels over time, comparing physical activity performance to other users and online coaching. A more affordable approach is self-reporting and multiple websites offer this service (e.g. *dailyburn.com, thecarrot.com* and *movescount.com*). The main advantage of automatic logging (with a step counter) is the ease for the user. However, manual logging may increase awareness and automatic logging may not correctly evaluate every activity, which can lead to a decrease in activity [16]. Another problem with current technology designed to persuade people to be more physically active is the price. The average selling price is 100 US dollars and for some systems there is an annual price of an additional 100 US dollars.

---

1 http://www.directlife.com
2 http://www.mytrak.com
3 http://www.fitbit.com
4 http://www.fitbug.com
5 http://www.bodybugg.com
6 http://www.zamzee.com

Exergaming

A new way of promoting physically active behavior is exergaming (a portmanteau of exercising and gaming). This term describes a new generation of video games that uses physical activity as input instead of more traditional input devices such as a joystick or gamepad. Examples are Dance Dance Revolution where the user dances on a mat, the PlayStation EyeToy and Microsoft Kinect with which users can play games using full-body motions, and Nintendo's Wii gaming system. Research-oriented examples can, for instance, be found in the work of Mueller [17]. An increased body movement in the game increases players' level of engagement, which makes exergaming a suitable platform for implementing persuasive technology [18].

For the Wii, there have been several games that specifically focus on sports and health, such as WiiFit and WiiSports. These games can be sufficiently physically intense to be considered MIPA [19] [20]. Furthermore, such games can be a gateway to exercising more in general [21].

The problem with these console-based exergames is that they require the user to buy expensive equipment and that they are limited to one location. There are also significant delays between deciding to play and actually playing (turn on the TV, turn on the game-device, load the game, etc.). The approach of this study, a smartphone game that requires players to engage in MIPA, does not suffer from these problems, but does retain the benefits of exergaming.

Smartphone Applications and Games

The latest generation of mobile phones, called smartphones, offers interesting new design opportunities because of the phones' sensors, connectivity and ubiquitousness. A plethora of smartphone games and applications are available for these phones. However, few focus on persuading people to become more physically active.

Most of the smartphone applications that focus on persuading people to become more physically active show exercise suggestions or log the physical activeness of the user. Some of the logging applications use the GPS- or accelerometer-sensor to perform automatic logging of running or walking. However, the sensors are not used for anything else.

Most of the smartphone games that use the phone's sensors to persuade people to be more active only make use of the GPS-sensor. Examples are "PicoPoke[7]" which requires users to walk around and take photos; "Tapcloud[8]", an augmented reality game; and "Goldwalker[9]" and "Mobile Adventure Walks[10]" which use the GPS-sensor to explore a virtual world. Only a handful of games use the phone's accelerometer for physically active gaming. An example is "Baseball Speed[11]", which measures how fast a user can throw. None of the games use the accelerometer to persuade people to become more physically active in general, though.

Summary

Persuasive technology has been used to get people to engage more in MIPA. Most of this persuasive technology has focused on all-day physical activity, using step

[7] http://gambit.mit.edu/loadgame/picopoke.php
[8] itunes.apple.com/nl/app/tapcloud/id424174431t=8
[9] http://itunes.apple.com/us/app/goldwalker/id372683234?mt=8
[10] http://www.mobileadventurewalks.com
[11] http://itunes.apple.com/us/app/baseball-speed/id321926476?mt=8

counters and self-reporting, and is expensive. Exergaming is a new and promising approach to persuade people to become more physically active. However, current console-based exergames are expensive, not easily portable, and it takes time to start them up. Smartphone-based exergames do not suffer from these limitations. Moreover, smartphones offer the same type of sensors as gaming consoles and have the added benefit of being both ubiquitous and connected. The resulting potential of smartphone exergames has so far not been fully utilized, though.

In the current study, a smartphone game was created that makes more explicit use of the advantages of smartphone exergames. As stated earlier, the goal was to examine whether such a game could make people engage more in MIPA. The next section discusses the design of the game.

2 Game Design

2.1 Concept Development

The basic concept is a game that is played with a smartphone and requires players to engage in MIPA to play it. A brainstorm was performed to generate ideas concerning more concrete implementations of this concept. The result of this brainstorm was a multitude of ideas for smartphone games that each requires players to perform a certain kind of physical movement. For instance, one idea entailed that players had to put their smartphone in a pocket of their pants and then rotate their waist as if they were hula hooping.

Next, the technical feasibility of building a smartphone game that can detect the various physical movements required by these games was investigated by logging the accelerometer data while playing. In addition, the games were tested out on ten potential users. That is, persons that fit in the target group (young adult office workers) were asked to imagine that the games were installed on their phone and make the movements required by the game. The users were provided by the list of fifteen games and could choose which one they like to play. For instance, they were told that the goal of one of the games was to “cycle" as fast as they can, so they were asked to sit down, put their phone in one of their socks and make cycling movements for a minute. After playing the game they could challenge two friends to beat their score. The next player could accept, deny or delay the challenge. Except for one participant who delayed the challenge the rest accepted and played the game.

Based on the results from these explorations, the following criteria for the final concept were determined:

- The required physical movement should be simple (e.g. a throwing motion). Otherwise, it will be too technically challenging to recognize the movement.
- The required physical movements should be as little out of the ordinary and attention-grabbing as possible, some of the games were perceived as embarrassing.
- The required physical movement should be sufficiently intense to be considered MIPA; some of the games did not demand a sufficient level of physical activity.
- The game should require users to hold their phone firmly; some of games could lead participants to accidentally drop their phones.

- The game should provide players with constant feedback on their performance. Such feedback was repeatedly asked for by user test participants.
- The game should have a competitive element. Participants clearly enjoyed competing with other participants and challenging them to beat their scores.
- The game should provide players with the opportunity to play against each other at the same time. Participants were more enthusiastic when playing against others at the same time than when taking turns to set a score.

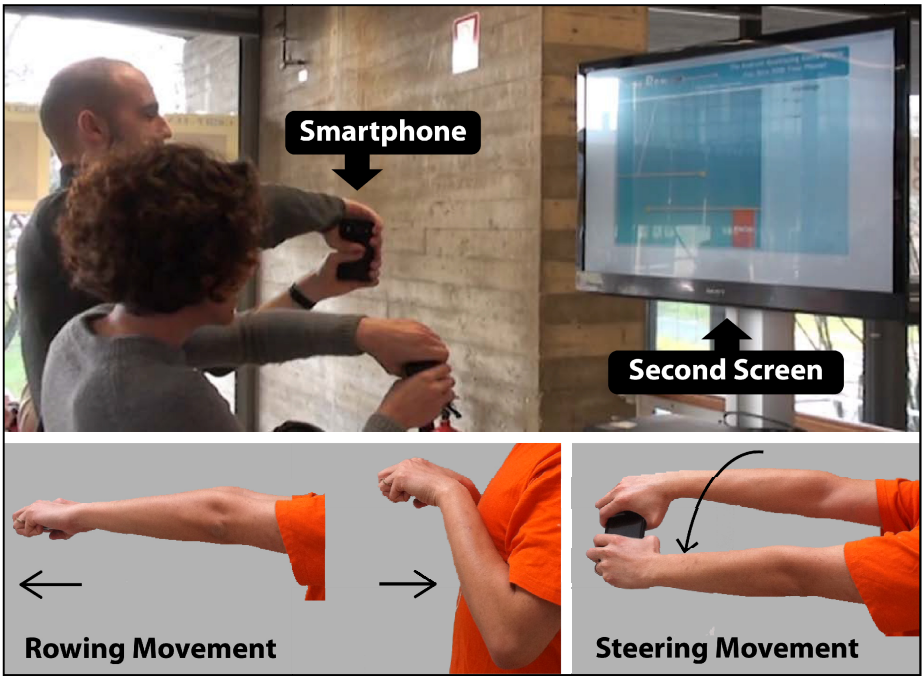


Fig. 1. Game setup and movements illustration

Based on these criteria the rowing-movement was selected (see Figure 1), this movement was very physically intense while it required users to keep their phone securely in their hands. It was not considered as embarrassing by the users. A second screen was added to provide feedback to the users and to allow players to 'row' against each other. The development of the final concept was split into two parallel tracks. One track consisted of developing the user interface by iteratively user-testing prototypes of the interface. These prototypes started out as paper prototypes and became progressively more high fidelity. The other track consisted of technical implementation of the game. When these two tracks were finished, they were combined to form the final game, which is described in the next section.

2.2 Game Description

In the final game, the players aim at 'rowing' their boats to the finishing line as quickly as possible by doing rowing and steering movements with their smartphones

(see Figure 1). Players compete with each other based on the time they spend on finishing the race. The racing result can be shared through a social network (Facebook) and the players can challenge each other with their highest scores. The players can check their overall ranking among all the players with a leaderboard.

The most important persuasive techniques implemented in this game are the competitive elements (face to face competition or comparing the racing results through Facebook) and the rewarding experience (winning the challenge or ranking high on the leaderboard), which are inspired by the persuasive strategies such as social proof, liking and praise [12][13].

Functional Description

The game consists of two parts, a smartphone-application and a website. To start, players go to the game's website (*www.yudugo.com*, see Figure 2 for a screenshot), with their computer's web browser. Here, they use their Android smartphone to scan a QR-code, which is to facilitate the installation of the application and joining the game, on the website. This QR-code links to the Android Marketplace where players can download and install an application, called "Phone Row". Next, players start up the application, enter a username and scan another QR-code on the website. As a result, they join a match and their username appears on the website. The website also shows a 30 second countdown timer. Within these 30 seconds, up to five other players have the opportunity to also join the match.

When the 30 seconds have passed, the website automatically goes to a new webpage with a ten second countdown timer. This webpage shows animated instructions concerning how to hold the smartphone, make rowing movements, and perform turns with their boats. On the smartphone, these instructions are also given with pictures and text.

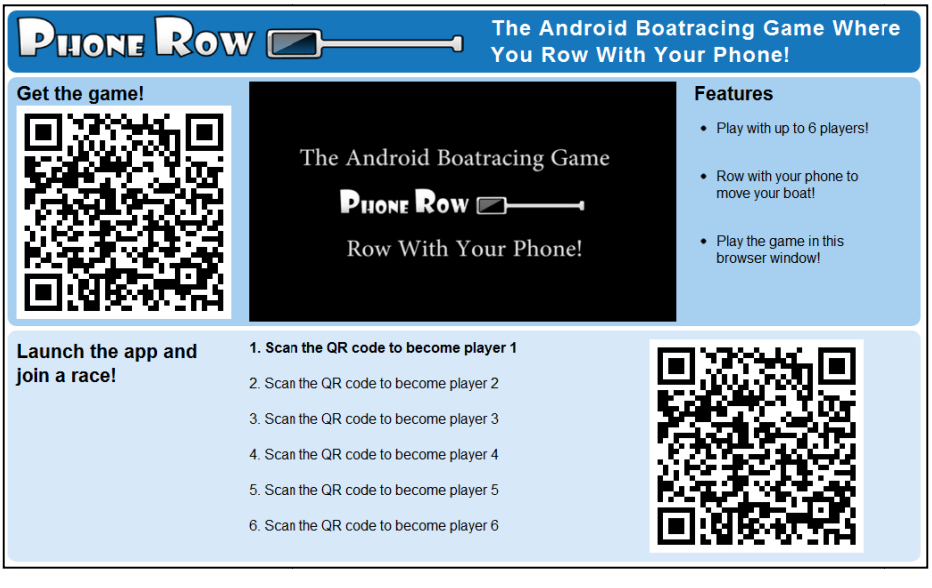


Fig. 2. The starting screen of the website

When the ten seconds have passed, the website automatically goes to a new webpage. On this webpage, the track and the boats of the players are displayed (see Figure 3 for a screenshot). The starting line is skewed to ensure that all players have to cover an equal distance. On the right, players can see which boat is theirs. The smartphone counts down from three and then the race starts. The players "row" with their phones to make their virtual boats move. The faster a player "rows", the faster the player's boat moves, although the speed of the rowing motions should be above a certain threshold to make the boat move at all.

When all the players have finished or five minutes have passed, the webpage automatically goes to the results page. On this page, players see in which position they finished, in what time they finished the track, whether they set a personal record, and what their position on the leaderboard is.

On the smartphone, the time in which a player has finished the track is also displayed. Moreover, the smartphone displays a Facebook button. When this button is pressed, players can login to their Facebook account and post their personal record on their own Facebook wall or the Facebook walls of friends and thereby challenge them to beat their personal record.

Technical Description

By using the phone's accelerometer the application measures a player's movement. Based on this movement, the player's position on the rowing track is changed. The application sends this to the website, which displays the player's current position.

To do this, the application needs to know which screen the player is currently viewing. When the application is first launched, it creates a specific identifier for itself (AppID). Every time the website is visited, it generates a new identifier for the

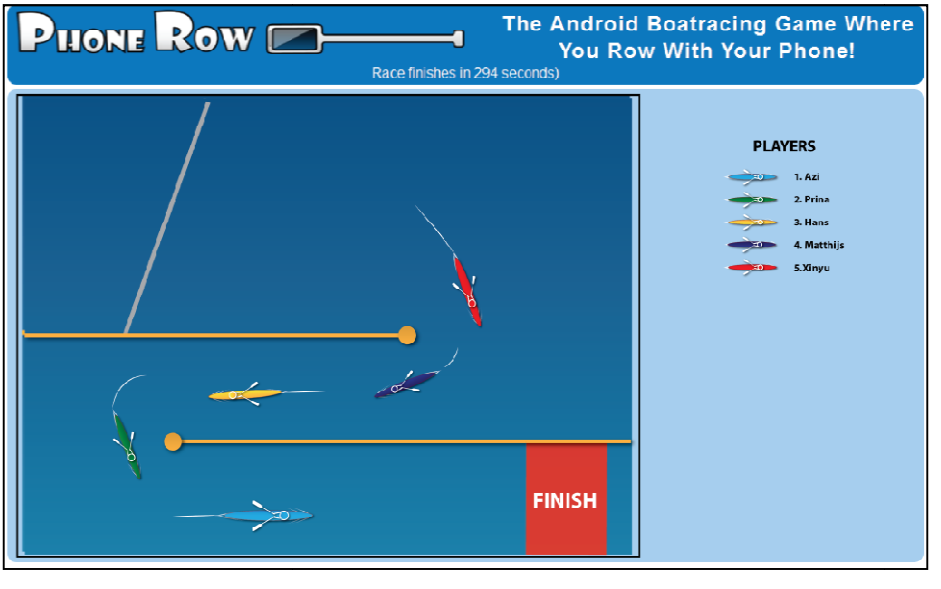


Fig. 3. The website shows the progress of the players

current computer screen it is displayed on (ScreenID). As mentioned earlier, to join a match, a player uses the application to scan a QR-code on the webiste. This QR-code contains the ScreenID. The application then sends its AppID, the ScreenID and the player's username to a script on the webserver.

After scanning, the application receives the start-time from the website. If this is the first scan, a 30 second countdown is initiated on the website. If it is not the first scan, the website sends a message back to the game stating how much time is left until the game starts.

Each application keeps track of a player's progress on the track and their time. As the time is measured by the game, it does not suffer from network latency and is accurate in the range of microseconds. Every 300 ms the player's current position is sent to a script on the webserver that stores it in the database. The website polls this database every 300 ms to display the progress of the player on the screen.

When a player crosses the finish-line, his or her total time is sent to a script on the webserver and is stored in the highscores list. This script sends back the finishing times of the other players. When all players have finished or after 5 minutes of time have past, the website shows the highscores as well. The application uses the ZXing-library[12] to scan QR-codes. The application sends messages to the website by sending an HTTP-GET request and receives messages in the JSON-format.

The website uses AJAX to update a webpage dynamically and uses the raphael[13] SVG-library for graphics. In the back-end, the website is programmed in PHP and MySQL. To create QR-codes, the website uses the Google Charts API[14].

3 Evaluation

The evaluation of the game focused on whether it became a habit for people participating in the study to play the game. That is, whether participants continued to play the game at least once a day over the four days following first time use. If this turned out to be the case, it could be suggested that persuasive technology, in the form of a smartphone game, can indeed help people engage more in MIPA.

The plan for getting a sufficiently large number of people to try the game consisted of two steps. The first step was to get a relatively small group of people to play the game. This step was executed by asking students from the Industrial Design department of the Eindhoven University of Technology, who own an Android smartphone, whether they were interested in trying out the game. Twenty students agreed to this.

The second step was to let these students spread the game among their friends. This step was supported by the Facebook functionalities of the game. As mentioned earlier, players have the possibility to post their personal records on their own Facebook wall or the Facebook walls of their friends and thereby challenge them to beat their personal record. The idea was that friends of people playing the game would see these wall posts and then become interested in trying out the game themselves.

[12] http://code.google.com/p/zxing/
[13] http://raphaeljs.com/
[14] http://code.google.com/apis/chart

The technical implementation of the game made it possible to anonymously track whether someone played the game more than once. However, this information could obviously not be used to understand why people chose to repeatedly play the game or not. As a result, the 20 students that agreed to play the game were asked, after they finished playing the game for the first time, what they did and did not like about the game. Also, usability problems that the students experienced when installing, starting up, and playing the game were noted down.

4 Results

4.1 Quantitative Data

58 people, including the 20 students, installed the game. Of these 58 people, 32 (55%) started playing a match and 15 (26%) completed a match. One person played more than one match. This person played the game a total of three times with approximately two hours between each match.

4.2 Qualitative Data

The twenty students that agreed to play the game had a wide variety of positive and negative remarks about the game. What follows is based on the remarks that were shared by most of the students. Most of the twenty students mentioned that playing the game felt like "real" physical exercise and that the concept of the game is novel and interesting. In addition, the interaction between the smartphone and the browser was deemed to be fun and innovative. Lastly, the student appreciated the competitive aspect. However, some students thought that it took too much physical effort to make the boats move, it was embarrassing to make the rowing movements and the game was not interesting enough to play more than once. Some users wondered if the movement was healthy or would lead to repetitive strain injury.

In addition to the comments, we observed several usability issues. It was not clear to the students when to look at the computer screen and when to look at the phone screen. As a result, important information, such as the instructions about how to "row", was sometimes not noticed by the students. It is not clear which physical movements should be performed to make the boat move forward or turn. Some students "rowed" with one hand instead of two and used small movements, which made the physical activity significantly less intense. Due to the implementation, it is not possible to make the boat turn and move forward at the same time, which was something that all participants wanted. The boats do not always move instantly after a rowing movement has been made, making it more difficult to learn the game. The game suffered from several software bugs; some of which resulted in the game crashing.

5 Discussion

5.1 Conclusion

The goal of the study was to create a game to help people engage more in MIPA. Thus, the game should be sufficiently physically intense and played on a regular

basis. Based on user feedback and observations, it can be concluded that the first requirement is met: playing the game requires MIPA. However, the second requirement is not met, because the current game concept was not implemented optimally. Although user input was actively sought throughout the design process, the current prototype did not provide players with a positive experience. In the current prototype feedback is given with a second screen, which was greatly appreciated. However, this introduced new HCI-problems, such as how to direct the users' attention to the required screen. Especially the confusion with regard to how to "row", and the fact that the boats did not move instantly after every rowing movement, seemed to contribute to a negative first impression of the game. The movement required by the game was physically intense, but might result in some form of repetitive strain injury. This risk was somewhat lessened as users have to stop moving to steer which gives a moment of relief. Also, the movement users make depends on the shape of phone and people play differently; not everybody makes the intended movement. In the current implementation, for instance, one-handed play gives an edge. From the experiments it was seen that users make small movements with the phone at first. Although users should be encouraged to make large movements for MIPA, these small movements should also result in an action in the game to make it easier for the user to learn the required movement by trial and error.

5.2 Future Work

To have an influence in the long run, users need to keep playing the games. Of the highest importance is solving the technological and usability problems. To improve the replay value, different games need to be added. The existing game also needs to be made more interesting by for instance adding more levels with increasing difficulty. Instead of using only the accelerometer, the touch screen, microphone and camera, can be used as well.

In the current approach, new players are attracted by approaching and challenging them through social networks. This did not have the effect that was expected. This might be because the correct audience was not reached. To reach the correct audience, advertisement is more suited. One way is to specifically target user groups by adapting the game to them or to use a current trend. For instance, rowing fraternities could be targeted for the rowing game and it is possible to make a "shake the banker"-game for the current financial crisis.

Currently, players are all in the same location and they use their own phone to control an object on a central screen. This concept can be expanded by allowing players to play over a distance, have games without a central screen or to have multiple central screens in one location. Finally, besides sharing scores to challenge other players, it could be possible for users to create their own motions and challenge other players with those.

This study shows an innovative way of exergaming, the use of the accelerometer in a mobile phone, extended with a second screen has not been explored before. User appreciation shows there is a great potential in the concept.

References

1. Norton, K., Norton, L., Sadgrove, D.: Position statement on physical activity and exercise intensity terminology. Journal of Science and Medicine in Sport 13, 496–502 (2010)
2. Woodcock, J., Franco, O.H., Orsini, N., Roberts, I.: Non-vigorous physical activity and all-cause mortality: systematic review and meta-analysis of cohort studies. International Journal of Epidemiology 40, 121–138 (2011)
3. Ainsworth, B., Haskell, W., Herrmann, S., Meckes, N., Bassett Jr., D., Tudor-Locke, C., Greer, J., Vezina, J., Whitt-Glover, M., Leon, A.: 2011 compendium of physical activities: A second update of codes and met values. Medicine & Science in Sports &Exercise 43, 1575–1581 (2011)
4. Wen, C.P., Wai, J.P.M., Tsai, M.K., Yang, Y.C., Cheng, T.Y.D.C., Lee, M., Chan, H.T., Tsao, C.K., Tsai, S.P., Wu, X.: Minimum amount of physical activity for reduced mortality and extended life expectancy: a prospective cohort study. The Lancet. 378, 1244–1253 (2011)
5. Van der Lucht, F., Polder, J.: Towards better health - The Dutch 2010 Public Health Status and Forecasts Report, `http://www.rivm.nl/bibliotheek/rapporten/270061011.pdf` (accessed December 12, 2011)
6. World Health Organization: Global recommendations on physical activity for health, `http://www.who.int/entity/dietphysicalactivity/publications/9789241599979/en/index.html` (accessed December 7, 2011)
7. U.S. Department of Health and Human Services: 2008 physical activity guidelines for Americans, `http://www.health.gov/paguidelines/pdf/paguide.pdf` (accessed December 7, 2011)
8. U.S. Department of Health and Human Services: Physical activity guidelines advisory committee report, `http://www.health.gov/paguidelines/Report/pdf/CommitteeReport.pdf` (accessed December 8, 2011)
9. Ku, P., Fox, K.R., McKenna, J., Peng, T.: Prevalence of leasure-time physical activity in Taiwanese adults: Results of four national surveys, 2000-2004. Preventive Medicine 43, 454–457 (2006)
10. IJsselsteijn, W.A., de Kort, Y.A.W., Midden, C., Eggen, B., van den Hoven, E.: Persuasive Technology for Human Well-Being: Setting the Scene. In: IJsselsteijn, W.A., de Kort, Y.A.W., Midden, C., Eggen, B., van den Hoven, E. (eds.) PERSUASIVE 2006. LNCS, vol. 3962, pp. 1–5. Springer, Heidelberg (2006)
11. Moere, A.V., Tomitsch, M., Hoinkis, M., Trefz, E., Johansen, S., Jones, A.: Comparative Feedback in the Street: Exposing Residential Energy Consumption on House Façades. In: Campos, P., Graham, N., Jorge, J., Nunes, N., Palanque, P., Winckler, M. (eds.) INTERACT 2011, Part I. LNCS, vol. 6946, pp. 470–488. Springer, Heidelberg (2011)
12. Cialdini, R.B., Goldstein, N.J.: Social Influence: Compliance and Conformity. Annual Review of Psychology 55, 591–621 (2004)
13. Fogg, B.J.: Persuasive Technology: Using Computers to Change What We Think and Do. Morgan Kaufmann Publishers, San Francisco (2003)
14. Kaptein, M., Lacroix, J., Saini, P.: Individual Differences in Persuadability in the Health Promotion Domain. In: Ploug, T., Hasle, P., Oinas-Kukkonen, H. (eds.) PERSUASIVE 2010. LNCS, vol. 6137, pp. 94–105. Springer, Heidelberg (2010)
15. Gasser, R., Brodbeck, D., Degen, M., Luthiger, J., Wyss, R., Reichlin, S.: Persuasiveness of a Mobile Lifestyle Coaching Application Using Social Facilitation. In: IJsselsteijn, W.A., de Kort, Y.A.W., Midden, C., Eggen, B., van den Hoven, E. (eds.) PERSUASIVE 2006. LNCS, vol. 3962, pp. 27–38. Springer, Heidelberg (2006)

16. Consolvo, S., Everitt, K., Smith, I., Landay, J.A.: Design Requirements for Technologies that Encourage Physical Activity. In: Proceedings of ACM CHI 2006 Conference on Human Factors in Computing Systems, pp. 457–466. ACM Press (2006)
17. Mueller, F., Stevens, G., Thorogood, A., O'Brien, S., Wulf, V.: Sports over a Distance. Journal of Personal and Ubiquitous Computing 11, 633–645 (2006)
18. Bianchi-Berthouze, N., Kim, W.W., Patel, D.: Does Body Movement Engage You More in Digital Game Play? and Why? In: Paiva, A.C.R., Prada, R., Picard, R.W. (eds.) ACII 2007. LNCS, vol. 4738, pp. 102–113. Springer, Heidelberg (2007)
19. Miyachi, M., Yamamoto, K., Ohkawara, K., Tanaka, S.: METs in Adults While Playing Active Video Games: A Metabolic Chamber Study. Medicine & Science in Sports and Exercise 42, 1149–1153 (2009)
20. Graves, L.E.F., Ridgers, N.D., Williams, K., Stratton, G., Atkinson, G., Cable, N.T.: The Physiological Cost and Enjoyment of Wii Fit in Adolescents, Young Adults, and Older Adults. Journal of Physical Activity and Health 7, 393–401 (2010)
21. Lieberman, D.A., Chamberlin, B., Medina, E., Franklin, B.A., Scanner, B.M., Vafiadis, D.K.: The Power of Play: Innovations in Getting Active Summit 2011 – A Science Panel Proceedings Report From the American Heart Association. Circulation 123, 2507–2516 (2011)

Developing Persuasive Technology for ASD Challenged Teenagers

Morten Aagaard and Peter Øhrstrøm

Department of Communication and Psychology, Aalborg University
Nyhavnsgade 14, 9000 Aalborg, Denmark
{morten,poe}@hum.aau.dk

Abstract. The HANDS project suggests the use of Mobile Persuasion in order to support teenagers with an autism diagnosis and normal or high IQ. The paper offers a description of the HANDS toolset and its potential. The HANDS toolset has been evaluated at four schools for teenagers with autism over a period of 7 months. The paper presents the main conclusions from this evaluation and some perspectives of the use of systems like HANDS are discussed.

1 Introduction

HANDS is an EU financed research project (June 2008 – October 2011). The name of the project stands for "Helping Autism-diagnosed teenagers Navigate and Develop Socially" (see [1]). 10 partners from 6 different countries have been involved in HANDS, and the project has been coordinated from Aalborg University, Denmark.

Autism spectrum disorders (ASD) – including, among others, Asperger's Syndrome – are developmental disorders of the human nervous system. They result in a lifelong condition (see [12]). Many individuals with ASD find it difficult to handle daily tasks, including confrontations with other people, and in many cases this leads to social marginalisation. Most individuals with autism require lifelong social, adequate psychological/pedagogical intervention.

The purpose of the HANDS project has been to improve the quality of life for teenagers with an autism diagnosis and normal or high IQ by providing an ICT toolset supporting them and their teachers in designing assistive and persuasive software tools, which can support the teenagers in their social development. The idea has been that each teenager in the project should be offered an individually designed toolbox implemented on his or her smart phone. The project has been based on the hypothesis that the use of such a HANDS toolbox can help the teenagers in improving their social skills and self-management skills, and thus ensure their social integration and independence. The HANDS toolset has been developed based on ideas from persuasive technology. The tools are implemented on smart phones, which are in communication with the HANDS server. The toolset makes it possible for a teacher at each of the partner schools to tailor special and individual tools for his students. These tools are supposed to help the users in their daily life. Two prototypes of the toolset

M. Bang and E.L. Ragnemalm (Eds.): PERSUASIVE 2012, LNCS 7284, pp. 67–78, 2012.

have been developed on a Microsoft platform. In addition to the two Microsoft prototypes the project partners have developed a prototype for an Android platform.

Many teenagers with autism can specify particular situations, which they find difficult to handle (see [11]). The elements of the HANDS toolset have been developed in order to make it possible for the teacher and the student in cooperation to design specific tools, which can support the student in dealing with the problems in such difficult situations.

Each of the prototypes of the HANDS toolset have been put in to practice and evaluated based on observations, interviews, logdata stored on the server and various tests carried out at the four partner schools in the HANDS project. A careful analysis of the test results can be found on the HANDS web-site.

In section 2 we shall outline the general design of the Hands systems and also discuss the use of ideas from persuasive technology in the systems including some of the ethical problems related to the systems. In section 3 the evaluation of the systems during 7 months from November 2010 to May 2011 will be presented. Section 4 will focus of the main conclusions from the HANDS project, and the final section will concentrate on future challenges in using persuasive technology in order to support marginalized teenagers with autism.

2 The Design and the Use of the HANDS Toolset

The HANDS toolset has been designed based on a study of the traditional practice at the schools for teenagers with autism. This means that the toolset is supposed to be used as an extension of the standard practice at the schools. The actual HANDS tools have been designed by the student and his or her teacher in a direct cooperation.

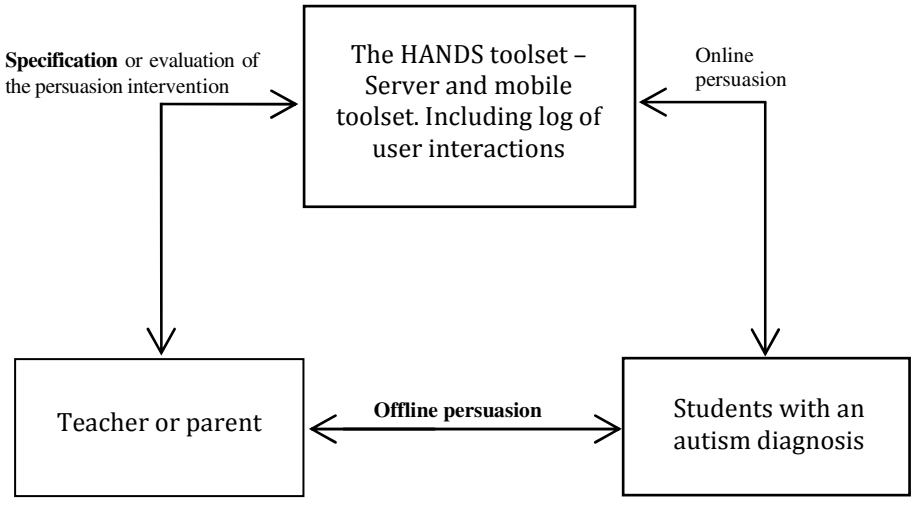


Fig. 1. Using the HANDS toolset available from the HANDS server the teacher and the teenager can co-operate in developing individually tailored tools, which may be helpful for the teenager in his daily life

The teachers are supposed to work closely together with their students in order to make individually tailored tools, which may function as a support tool for the teenagers in their daily life in school and elsewhere. The toolset is made available on the HANDS server and all interactions with the HANDS tools will be stored (logged) on the HANDS server.

The HANDS toolset includes the following functionalities:

1. The Handy Interactive Persuasive Diary (HIPD).
2. The Simple-Safe-Success Instructor (SSSI).
3. The Personal Trainer (PT)
4. The Individualiser (TIn)
5. The Sharing Point (SPo)
6. The Credibility-o-Meter (CoMe)

HIPD is designed to facilitate the temporal organization of the activities, which are important for the teenagers. The interactions are calendar functions - such as seeing, modifying, adding appointments, seeing the day plan or week plan and using prompts. This functionality can support the student in managing his or her daily schedule better. (See [10]). At certain times the system may send a prompt to the student. This may be combined with a question which the student is going to answer by picking one option from a short list of possible answers presented on the smart phone.

SSSI and PT offer strategies for dealing with the individual teenager's problems. A specially prepared SSSI is designed to be used by the student when facing a specific problem which the student sees as difficult to handle, whereas a specially designed PT is supposed to be used by the student in order to prepare himself or herself for dealing with the difficult situation in question. In the present context, however, there is no need to distinguish between SSSI and PT. In the following we shall treat these two functionalities as one using the term PT. Obviously, the PTs can be designed to fulfil a variety of pedagogical purposes. The choice of the actual design will depend on which situations the student in question finds difficult. Each PT contains a number of steps corresponding to the strategy which the teacher advices the student to follow in order to deal with the difficult situation in question. As with HIPD the actual use of the specially designed PT can be stimulated by the use of various kinds of rewards. For this purpose, there is in fact a functionality for giving reward points built into the system.

TiN makes it possible for the teacher and the teenager in co-operation to tailor the tools in an individual manner. The aim is various kinds of individualization in terms of audio and visual skin. This functionality is essential if the tools are going to be accepted and actually used by the teenagers.

In Prototype 2 of the HANDS toolset the SPo functionality is only relevant for the teachers at the partner school who, having used this functionality, can share their experiences from the work with the toolset. Later an attempt at introducing social media to teenagers with autism in a safe manner was added to the SPo (see [1]).

CoMe is based on the logdata and other data stored on the HANDS server. It offers various ways for the teacher and for the researchers to follow the activities of the teenager in order to learn more about his or her interaction with the tools – the ultimate goal being a possibility of measuring how credible the teenager finds the tool.

The HANDS tools are designed to be persuasive. This means that the use of the tools should lead to a changed behaviour as compared with what would otherwise be expected. In many cases, the PTs may be closely related to the HIPD. This is at least true when it is possible to predict the difficult situation which has given rise to the PT. In such cases the PTs can simply be scheduled as an integrated part of the HIPD. In other cases the difficult situations which the PTs are designed to deal with cannot be predicted, but they may pop up relatively unexpected. In any case, the timing issue is very important. It is well known that a PT should be employed at the right time, the so-called Kairos (see [2]), in order to be maximally effective.

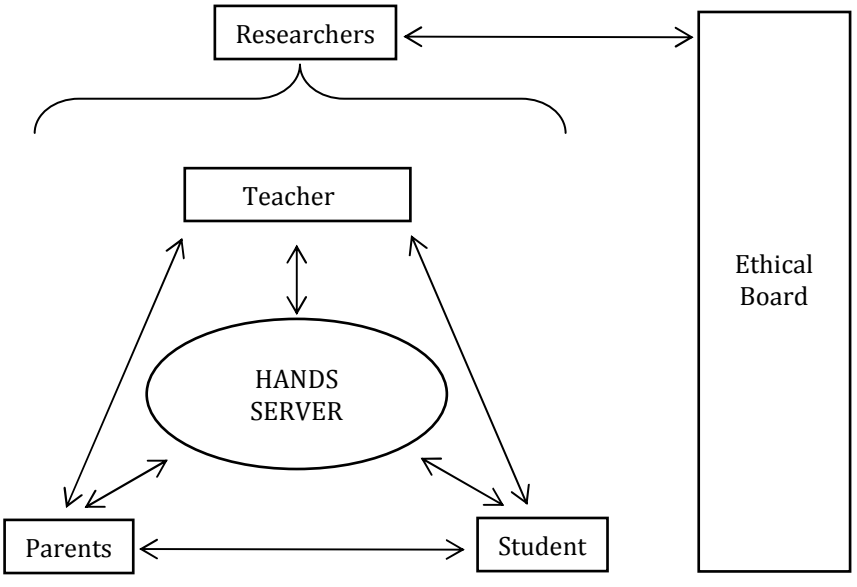


Fig. 2. The communicational setup in HANDS. Experiments etc. have had to be approved by the Ethical Board before they were carried out.

Clearly, this setup may give rise to many ethical questions. First of all, the HANDS procedures include a lot of surveillance in the sense that all activities involving direct use of the HANDS tools will be logged on the HANDS server. It may be discussed whether such an intensive use of surveillance may in itself be problematic (see [7]). The data on the server may be accessed by the student himself, by the teachers and the researchers, and to some extent also by the parents. This clearly means that there is at least a need for ethical discussions regarding the use of and the access to such person-sensitive data. Which data should be treated as private? The ethical problems with person-sensitive stored data are well known and have to be handled carefully taking the actual setting and context into account (see [13]). In order to deal with the ethical questions related to the HANDS project in a qualified manner, an Ethical Board has been established as an important part of the research setup. All experiments and pilot studies in HANDS project have had to be approved by the Ethical Board before they

could be carried out. The members of the Ethical Board have been independent in the sense that they did not participate either in the research at the schools, or in the testing procedures. However, the members of the Ethical Board have been invited to take part in the discussions at the general meetings in HANDS in order to strengthen the ethical and value-oriented dimensions of the work carried out within the HANDS project.

3 Evaluation of the HANDS Toolset

In analyzing the data, our focus was on a series of cases, defining cases as a teacher-student dyad. These cases have been studied as individual cases and the perspectives have been carefully discussed within the HANDS project (see [1], [8], and [9]). In addition, it has been attempted to obtain a statistical overview of the logdata stored on the server. In general, it is, however, difficult to do relevant population statistical analysis of data concerning individuals with autism. The reason is that they, as already mentioned, are very different. A general statistics will typically contain a variety that ranges from great failures and great successes. Both initiatives taken by the student and interactions otherwise occurring on the mobile itself will be logged on the server. Clearly, the server data in question contain important information regarding the user behaviour.

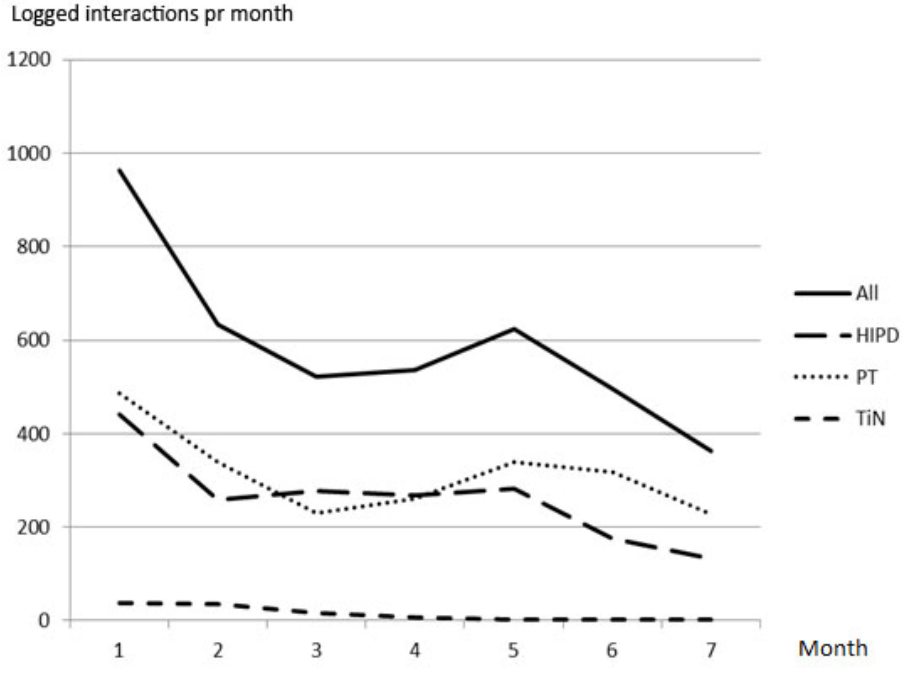


Fig. 3. Overall log statistics of Prototype 2 use (from November 2010 to May 2011) by the students at the HANDS partner schools. The figure includes statistics for three functionalities.

The above diagram illustrates the use of the HANDS tools (HIPD, PT and TiN) over a period of about 7 months from November, 2010, to May, 2011. During this period only HIPD, PT (SSSI) and TiN have been available for the students. In total, data from 20 students have been included in this evaluation. The diagram summarizes the log data corresponding to the use of these three elements of the HANDS toolset.

The rather high activity during the first month of the period should be seen as a result on the introduction of the software to the students and the teachers. After this introduction, the users are left to incorporate the toolset in their daily practice. From month 2 the activity is slightly decreasing, but later it is increasing, which is consistent with the assumption that users have to learn to apply the toolset and to appreciate its potential. Finally, the use of the activity is decreasing towards the summer period.

During the evaluation period a reward system has been running. By interacting with the system in certain desirable ways the student could earn reward points which could be used in order to obtain certain benefits at the schools. More than half of the students actually used the reward system during the period. The teachers found this option useful in the educational context.

The log results of the three functionalities are very much as expected. Using Calendar functions can be an everyday activity that generates many log data. The TiN is only used occasionally, when the student finds it is appropriate. Obviously, it is a functionality that does not create much interaction and therefore not much log data.
Most of the students did not use prompts very much during the evaluation period. However, 3 of the students became “heavy” users of prompts (75% of all the prompt interaction relates to them). The teachers have found this option useful in several educational contexts.

The Personal Trainer (PT) is a focused pedagogical intervention which is usually limited to a time interval or to certain types of activity. The construction of PTs will in general be more demanding and complicated than the practical use of the HIPD. However, in many cases it will be possible to integrate the PTs in the context of the HIPD. The use of PT calls for much attention and creativity from the student and his or her teacher. In the construction they have to concentrate on the situations which the student finds difficult to handle. The following is a selection of the PTs which have been designed by the users and applied during the evaluation period:

- Getting Stressed RED: What to do when you get stressed and need to leave the room
- Getting Stressed AMBER: What to do when you are getting stressed but feel you can stay in the room
- Travelling on a bus
- Making toast
- Out on my own and getting lost: What to do if you are worried when alone
- Using the phone
- What to do when you are anxious
- Talking to people: What you can talk about
- Steps to order a pizza
- How to have a drink at a café

- Blood test and insulin
- How to shop at the supermarket
- Good behaviour when losing at games
- Changing position in the classroom
- On food and nutrition: What to eat when…

It is obvious that all the PTs listed above are persuasive in the sense that each of them is aiming at a certain behaviour which is advisable for the student. In all cases the students have given their consent to putting the PTs into operation on their smart phones, since they have understood that this changed behaviour will help them to function better socially.

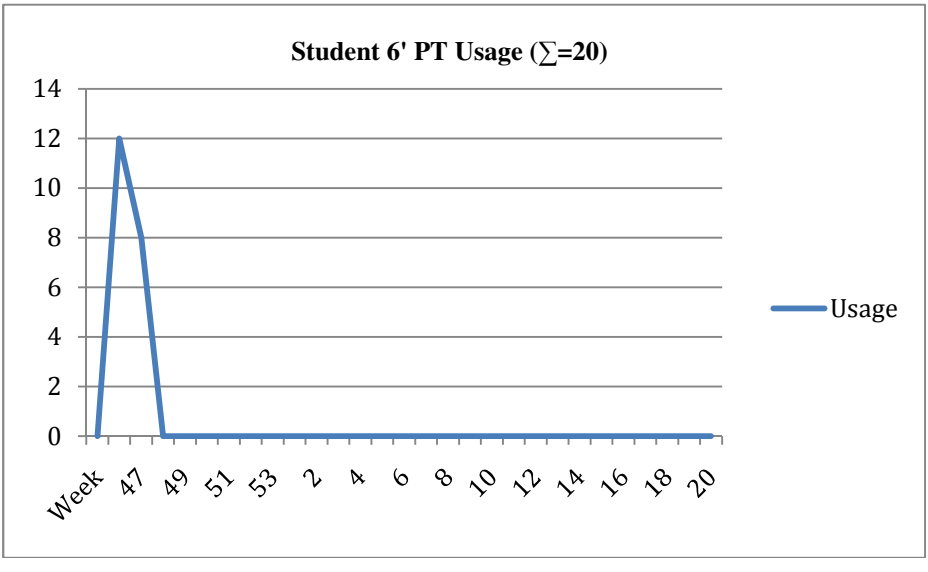


Fig. 4. Student 6' use of a certain PT during the evaluation period. For privacy reasons the actual content of this PT should not be published.

The logdata show that the "lifetimes" of the PTs may differ a lot. In some cases the PT is only used in a short period after it has been put into function. One possible explanation may of course be that the student does not find the PT credible, i.e., that he does not believe that it can be of any significant help to him. However, this is not the only possible interpretation of a diagram like the one in Fig. 4. Another possible explanation would be that the difficult situation in question only occurred for a short period. Alternatively, it may be that the student during two weeks of use of the PT became confident that he would be able to handle the difficult situations in question by himself without help from the HANDS tools. However, the teacher will, knowing the student and having access to the logdata, be able to determine which of these possible explanations will be the correct one. In fact, on the basis of the experiences from the evaluation period the teachers have reported that the access to the logdata is a valuable resource in the direct interaction with the student.

In other cases the logdata from the usage of a PT are as shown in Fig. 5. Here the PT is used over a long period. It looks as if the difficult situation for which the PT has been designed has occurred again and again during the evaluation period. And since student 2 has actually been using the tools repeatedly, it seems that he finds it credible and helpful for him. On the other hand, there may of course still be occurrences of the difficult situations in question at which the student has decided not to open the PT on his smart phone or simply forgotten to do so. If the teacher can obtain information concerning such occurrences, it can certainly be important in the direct interaction with the student. One other question which the teacher might need to consider has to do with the challenges caused by the repeatedly use of the tool: Will the student depend on having access to the PT whenever he is facing the difficult situation in question? Will the student become addicted to the tools? If this is the case, the teacher and student have to consider the various options carefully in order to find the best strategy.

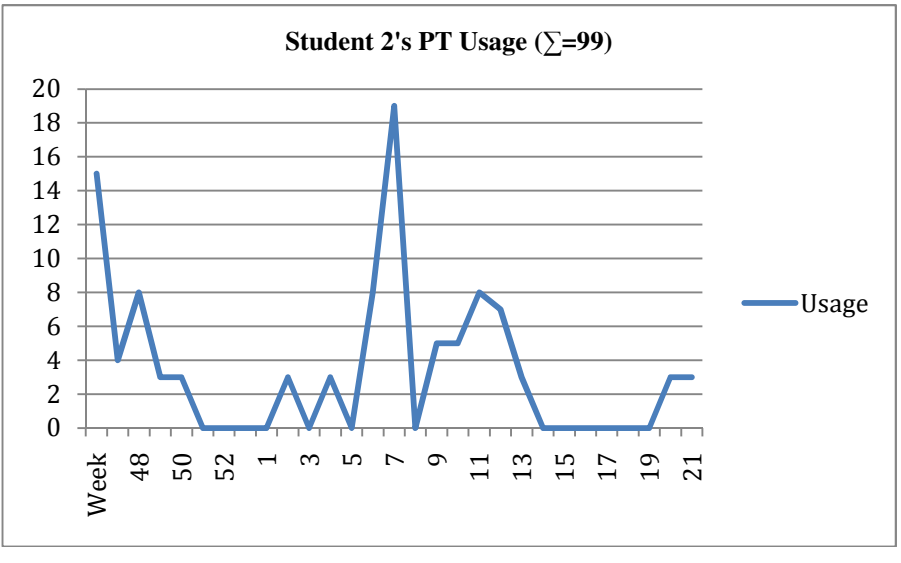


Fig. 5. Student 2's use of a certain PT during the evaluation period. For privacy reasons the actual content of this PT should not be published.

4 Some Main Conclusions from the Evaluation of the HANDS Toolset

The evaluation of the HANDS toolset should obviously be compared with the traditional methods based on paper and pencil. The main question is whether the potential of the methods based on the toolset is higher that the potential of the traditional methods. The evaluation gives evidence for the conclusion the HANDS-assisted interventions *can be significantly more effective* than traditional types of intervention and support. Clearly everything that can be done using the traditional methods based on paper and pencil, can also be done with HANDS-assisted

interventions. In addition, the use of the HANDS toolset offers extra possibilities in terms of interactivity with the tools and in terms of the log data stored on the server. In particular, the teachers at the partner schools have found that the use of the log data from the server can improve the effectiveness of the interaction with the student significantly. The tradeoff is that introducing and using tools such as the HANDS toolset requires that the teachers are capable of designing relevant and useful PTs together their students.

The role of the teacher in HANDS-assisted interventions

It is assumed that the decision to use the HANDS toolset is made on the basis of careful consideration of the individual user's specific support needs, and the actual focus and content of the intervention is set and designed on the basis of such considerations and a professional understanding of principles of psycho-educational intervention and support in ASD. In order to carry out such HANDS-assisted interventions it is of course essential that the teachers are able to master the HANDS toolset. This obviously requires some ICT skills. Within the HANDS project it has been underestimated what it takes to fulfil this condition. In order to benefit maximally from the potential of the HANDS toolset it is clearly important to set up relevant courses for the teachers who are going to develop and tailor the actual HANDS tools in cooperation with their students.

The need for a clear identification of the difficulties relevant for the intervention

The evidence indicates that in order to obtain an effective intervention the teachers and the students should clearly identify the difficulties they want to work with. The evidence indicates that it is very important that the teachers and the students focus clearly on identifying clear goals when working on behavioural changes in HANDS. It is evident that the accurate identification of the teenagers' needs by the teachers, followed by the formulation of well-bounded and clearly defined interventions on HANDS, are likely to lead to higher levels of engagement and response to behavioural interventions.

Where an accurate identification of the difficulties is undertaken, it needs then to be expressed in terms of clearly formulated interventions using HANDS tools. Thus some interventions will, by their nature and the functional ability of HANDS, lend themselves to clear formulation in HANDS, and some will not. The evaluation indicates that teacher awareness of such distinctions is an important factor in mediating the appropriateness of interventions developed on HANDS, and subsequent engagement with the intervention by the teenagers. The teacher expertise and teacher guidance should if possible be formulated in terms of so-called "tunnels" (see [4] and [14]). In several cases, the teachers have reported that the use of this technique is highly effective. In particular, this is the case when the behavioural interventions are focused on locations outside of the school domain.

There is evidence that of the potential ability of teachers to successfully and appropriately identify needs that can in many cases be formulated as interventions on HANDS, which may be very helpful for the students in their daily life. Hence, it is concluded that the positioning of social and life skills development, particularly in terms of the overall school curriculum and approach, is a significant factor in order for the use of solutions like the HANDS toolset to be effective.

Persuasion and credibility; Human vs HANDS interventions
The literature on persuasive technology (see [4]) indicates that in some instances computers can be regarded as more credible sources for persuasive messages than human actors (see [5]). The evaluation indicates that some teenagers with ASD have a preference for receiving persuasive interventions from a mobile device in comparison to their teacher in some contexts. In some cases, this is because children do not perceive the mobile device as having the overbearing or "nagging" quality of their teacher. In that case, a greater focus on the identification of their own needs can be regarded as an expression of their autonomy. Similarly, the teenagers may regard the device as a removal from the perceived authority of their teacher, and thus may feel that their sense of autonomy is better preserved when they receive messages from the mobile device as compared to their teacher. In other cases, there is a mode preference for the reception of sequential instructions, as on the mobile device, particularly with the PT function, the student can control the flow of information.

Credibility and the need for a Mobile Marriage
It is important that the teenager develops a positive relationship with the device, based on repeated interactions over a period of time (a so-called Mobile Marriage). These repeated interactions will typically involve other mobile phone functions such as the use of SMS, Internet, Social Media etc. There is evidence that such a mobile marriage will increase the likelihood that the user will respond positively to persuasive messages delivered via a specific persuasive mobile system such as HANDS. In fact, the data provides strong evidence to support the contention that mobile marriage is a significant factor in mediating the student's engagement with the HANDS tools and their response to behavioural interventions. If they don't view the phone itself as having credibility, they are less likely to want to engage with it, and in the perception of students, parents and teachers, less likely to respond positively to persuasive messages from HANDS.

Communication with parents
The evaluation indicates that there is a need for a greater liaison with the students' family when using the HANDS software. Parental input can facilitate out of school use, thus maximizing the portability potential of HANDS, and the potential to exploit fully opportunities for social and life skills development outside of the school.

Practically, it is important that teachers communicate with parents when planning interventions timed to activate when the teenager is at home with their parents. Furthermore, parents can contribute with vital pieces of information about overall social and life skills which teachers may otherwise be unaware of. There is evidence that where there has been effective liaison with teachers, this has been associated with the successful development of out of school interventions. It has been documented that the introduction of the HANDS tools can facilitate an increased communication between home and school.

Technical issues
The technical stability of the software and also the perceived reliability by teachers and students clearly affect how credible the device is thought to be. This is a particularly important factor for use outside rather than inside the school, as within

the school there are alternative support structures more readily available such as pen and paper forms of support as well as staff such as teaching assistants, who can provide on the spot support if the technology fails or does not quite match the needs of a particular situation. If the teenager is using the phone independently, an adult may not be there to support them. In such circumstances, the reliability of the HANDS tools has more relevance.

There is evidence that platform flexibility is a significant factor mediating, in particular, the potential use of HANDS in out of school settings. Specifically, the evaluation indicates that in order for more out of school use to take place, greater flexibility is also needed between the mobile device and the web based CoMe programme.

5 The Use of Persuasive Technology in Order to Support Teenagers with Autism – Some Future Challenges

The motivation behind the HANDS project is the fact that the behavioural and cognitive characteristics of people with ASD put them into a high risk of social marginalization. As argued by Miklos Gyori and others the impairment in reciprocal social behaviours may direct prevent the affected individual from participating in everyday social interactions, and the oddness of social behaviours often leads to rejection by others in the social environment, and also anxiety in the affected person (see [6, p.19]). However, teachers at schools for teenagers with autism have over the years developed rather effective methods in order to support teenagers with ADS who would otherwise be socially marginalized.

Much of the evidence provided in the HANDS project is based on qualitative methods. It would be interesting also to carry out a large (quantitative) control experiment in order to compare the HANDS methods with the traditional methods (based on pen and paper). However, the HANDS project has provided evidence for the fact that teachers can support teenagers with autism even better by transforming these traditional methods into the context of persuasive technology. Persuasive mobile technology may essentially support all the traditional methods based on paper and pencil. In addition, the HANDS project has demonstrated that the tailored tools on the smart phone can be more effective than traditional methods by being interactive. Furthermore, in HANDS the actual use of the tools will be logged on the HANDS server, which gives the teacher a valuable input for following-up in the personal interaction with the student.

Some teachers may see the HANDS-assisted intervention methods as more complicated than the traditional methods. It is obvious that potential of the HANDS toolset can only be exploited fully if the teachers at the schools for teenagers with autism are properly introduced to the toolset to the extent that they become able to master the systems. This is in fact what motivates the HANDS Open project (see [3]) which may be seen as a continuation of the HANDS project.

Acknowledgements. We are grateful to the entire HANDS team as this paper is in fact based on the results of the HANDS project as a whole. In particular, we want to thank Joseph H. Mintz, London South Bank University, as well as Jaap Erasmus and

James Collett, Helen Allison School, The National Autistic Society, for useful comments on earlier versions of this paper.

References

1. The Hands Web-site, http://hands-project.eu/
2. Aagaard, M., Øhrstrøm, P., Moltsen, L.: It might be Kairos. In: Persuasive 2008, Oulo, Finland (2008)
3. Aagaard, M, Gyori, M., Kanizsai-Nagy, I., Stefanik, K., Aaen, M., Mintz, J.: HANDS – Plan for use and dissemination. HANDS Project deliverable D1.3.1B (2011)
4. Fogg, B.J.: Persuasive Technology - Using computers to change what we think and do. Morgan Kaufmann Publishers, San Francisco (2003)
5. Gerdes, A., Øhrstrøm, P.: The Role of Credibility In the Design of Mobile Solutions to Enhance the Social Skill-Set of Teenagers Diagnosed with Autism. Journal of Information, Communication & Ethics in Society (2011)
6. Gyori, M., Kanizsai-Nagy, I., Stefanik, K., Vígh, K., Öszi, T., Balázs, A., Stefanics, G.: Report on initial cognitive psychology requirements on software design and content design & content. HANDS Poject deliverable D2.2.1 (2008)
7. Leth Jespersen, J., Albrechtslund, A., Øhrstrøm, P., Hasle, P., Albretsen, J.: Surveillance, Persuasion, and Panopticon. In: de Kort, Y.A.W., IJsselsteijn, W.A., Midden, C., Eggen, B., Fogg, B.J. (eds.) PERSUASIVE 2007. LNCS, vol. 4744, pp. 109–120. Springer, Heidelberg (2007)
8. Mintz, J. , Aagaard, M.: The Application of Persuasive Technology to Educational Settings. Educational Technology Research and Development, pp. 1–17 (January 26, 2012)
9. Mintz, J., Branch, C., March, C., Lerman, S.: Key factors mediating the use of a mobile technology tool designed to develop social and life skills in children with Autistic Spectrum Disorders. Computers & Education 58(1), 53–62 (2012)
10. Ranfelt, A.M., Wigram, T., Øhrstrøm, P.: Towards a Handy Interactive Persuasive Diary for Teenagers with a Diagnosis of Autism. In: Proceedings of the 4th International Conference on Persuasive Technology. ACM International Conference Proceeding Series, vol. 350, p. 279s. Association for Computing Machinery. ACM (2009), http://portal.acm.org/citation.cfm?doid=1541948.1541953
11. Schärfe, H., Øhrstrøm, P., and Gyori, M.: A Conceptual Analysis of Difficult Situations - developing systems for teenagers with ASD. In: International Conference on Conceptual Structures (ICCS 2009). CEUR Workshop Proceedings, vol. 483 (2009), http://ftp1.de.freebsd.org/pub/publications/CEUR-WS/
12. Volkmar, F.R., Lord, C., Bailey, A., Schultz, R.T., Klin, A.: Autism and pervasive developmental disorders. Journal of Child Psychology and Psychiatry 45(1), 135–170 (2004)
13. Øhrstrøm, P., Dyhrberg, J.: Ethical problems inherent in psychological research based on internet communication as stored information. Theoretical Medicine and Bioethics 28(3), 221–241 (2007)
14. Øhrstrøm, P.: Helping Autism-diagnosed Teenagers Navigate and Develop Socially Using E-learning Based on Mobile Persuasion. International Review of Research in Open and Distance Learning 12(4), 54–71 (2011)

Talk to Act: How Internet Use Empowers Users to Participate in Collective Actions Offline

Sandy Schumann[1], Olivier Klein[2], and Karen Douglas[3]

[1] Université Libre de Bruxelles & F.R.S.-FNRS
[2] Université Libre de Bruxelles
[3] University of Kent at Canterbury

Abstract. This study examines how Internet use can empower users to carry out collective actions for an environmentalism movement organization. More precisely, we focused on the impact of online interactivity, i.e., the fact that users can share content online and receive feedback on it from others. The participatory Internet fulfills thereby two preconditions of a sense of psychological empowerment: a) receiving information about the goals and performance of an organization and b) experiencing an effective reward system. Using an experimental design, our results showed that users' sense of empowerment was indeed increased by online interactivity. Higher sense of empowerment led to stronger willingness to participate in a panel discussion and demonstration for the environmentalist organization. In addition, when users were identifiable with their name and photo as compared to being anonymous while making their contributions, the likelihood to get engaged was higher, mediated by an increased sense of empowerment. The importance of intra-individual processes when studying the impact of Internet use on behavior is discussed, as well as the role of identifiability online.

Keywords: psychological empowerment, collective actions, social media, anonymity.

1 Introduction

The Internet is often viewed as an individuating technology [1]. But previous research also showed [2] that Internet use can facilitate participation in civic and political collective actions, where individuals are representing a group[1] they belong to and aim to improve its status, power, or influence [3]. However, there are only few systematic analyses of the processes that underlie the relationship between Internet use and collective action participation. The goal of this article is to contribute to this research. We do so by considering the affordances of the participatory Internet, namely bidirectional communication processes, i.e., online interactivity [4], which allow users to be both recipients and senders of content. In line with previous work, we show that online interactivity leads to greater willingness to participate in offline collective actions, through an increased sense of psychological empowerment.

[1] In the article, we use the terms group and movement organization interchangeably.

M. Bang and E.L. Ragnemalm (Eds.): PERSUASIVE 2012, LNCS 7284, pp. 79–89, 2012.

Our results provide important insight about why isolated group members act on behalf of their group that is only psychologically, but not physically, co-present. In addition, implications for the theoretical frameworks that guide the study of the Internet's impact on (offline) behavior can be derived. We advocate a perspective that focuses on intra-individual processes, and extends the current psychological approach, which examines largely inter-individual differences [5]. Previous research relied often on cross-sectional designs. Applying a longitudinal one allows causal conclusions for design elements that can help movement organizations to use the Internet to attract supporters.

2 The Effects of Internet Use on Collective Actions

When studying the effect of Internet use on participation in collective actions, Internet-supported and Internet-based actions are distinguished [6]. The earlier refer to offline actions such as demonstrations that are facilitated through online discussion groups or Email for example; the latter include cyber sit-ins, hacking or signing online petitions and are solely carried out online [6]. The focus of this article is on Internet supported actions. Previous studies showed that receiving information online (e.g., political news on web sites and video channels, information shared in online discussions) did increase the level of participation in civic and political collective actions offline [2], [7]-[10]. Fisher and Boekkooi [11] analyzed for example how participants of a day of climate actions were mobilized through the Internet. They concluded that the information provided online encouraged especially disconnected, but motivated, supporters by making them aware of the events and allowing them to coordinate their participation easily.

The general interpretation of the results follows an instrumental approach. It hypothesizes a relationship between the costs of receiving information and engagement [5]: Internet use should reduce the costs of knowing the goals and activities of a movement organization, which in return should increase the likelihood of participating in collective actions. In addition, Internet use is proposed to reduce the costs of having access to a greater variety of insights. Gil de Zúniga and Valenzuela [12] showed that Internet use increased especially the number of weak ties. Compared to strong ties, such as family and friends, which often provide redundant information [13], weak ties are more likely to share what is new and not yet known. Indeed, the number of weak ties mediated the effect of online conversation on civic engagement, which was stronger than that of offline conversations [12].

3 Online Interactivity and Collective Actions

This instrumental approach does not take the affordances of the participatory Internet into consideration, and views Internet use as a top-down communication process [14]. However, Internet use is defined by online interactivity that allows users also to contribute content and to receive feedback on it from others (e.g., supporters or representatives of movement organization) through comments or "likes". Defining Internet use in this way, we propose that it can facilitate one's willingness to participate in collective actions by increasing the user's sense of psychological empowerment.

Psychological empowerment is defined as the motivational concept of self-efficacy [15] or an intrinsic task motivation [16]. It is multidimensional and encompasses meaningfulness (i.e., the goal of the task is valuable in relation to one's own ideals), competence (i.e., one believes in one's own capability to be skillful), self-determination (i.e., one feels like having the choice in regulating the action) and impact (i.e., one feels like having influence on reaching a higher level goal) [15], [17]. Higher empowerment was shown to affect task performance, innovative and organizational citizenship behavior [18].

Contextual and individual characteristics were postulated to influence one's sense of empowerment. Spreitzer [15] summarized as the earlier information and rewards and as the latter self-esteem and one's locus of control (i.e., the feeling that either oneself or external forces determine what happens). Our research question explored how online interactivity can affect one's willingness to participate in collective actions that take place offline. Hence, contextual antecedents that increase the sense of psychological empowerment have to be reviewed. Kanter's [19] early work stated that a context where information about the mission and performance of an organization is provided is crucial to enable individuals to act in line with the organizations' goals, creating a sense of purpose. It further allows the reinforcement of a sense of competence and influences decisions to uphold one's engagement [15]. Through their online presences, movement organizations can share these insights and increase thereby the users' sense of empowerment. Communicating the group's mission and activities allows users to infer norms and define what is perceived as prototypical and valuable [20]. This is deeply interconnected with the second antecedent of an increased sense of empowerment: a reward or incentive system [15]. Individual contributions can be rewarded when in line with the group's position and sanctioned if that is not the case. Of course, a reward system is only effective if everyone's input is noticed by others. Only then can individuals be positively evaluated by their group and recognize in return how their actions are part of the efforts to achieve the organizational goals [15]. Ultimately, when users can share their opinions and stories online and have the option to receive feedback on them, their sense of empowerment should be even more heightened.

If one's sense of empowerment depends partially on rewards, users that act in line with the group's norm should be more rewarded by the group and feel hence more empowered. Providing personal information about users online should influence their motivation to act as a prototypical group member [21], i.e., their need for positive self presentation to the group [22]. Indeed, research in the framework of the social identity model of de-individuation effects [23] showed that being identifiable to one's group elicited the need of presenting oneself in line with the group norms [22]. This is referred to as the strategic effect [23] of identifiability online. It elicits normative behavior as well as a sense of accountability for one's actions [22].

Based on this argument, we postulate the following hypotheses:

Hypothesis 1: Online interactivity increases the users' sense of psychological empowerment.

Hypothesis 2: Higher psychological empowerment leads to a higher willingness to participate in collective actions.

Hypothesis 3: Being identifiable as compared to anonymous when generating content leads to a higher willingness to participate in collective actions, mediated through a higher sense of empowerment.

4 Method

4.1 Design

We tested these hypotheses with a mixed design. Participants were randomly assigned to one of two experimental conditions, in which they were either identifiable with their full name and photo, or anonymous through a standardized user name (e.g., ParticipantA). Data was collected at three measuring points (see 4.3).

4.2 Sample

The $N = 84$ participants were undergraduate psychology students at the University of Kent and received class credit for their participation. To avoid possible gender effects in the identifiable, but not in the anonymous condition, all participants were female. They were on average $M = 19.77$ ($SD = 3.79$) years old.

4.3 Procedure

The study was framed as being concerned with understanding how people feel about protecting the environment. A group of three or four participants was tested at the same time. Participants did not know who the other participants were, as they were seated in separate rooms and great care was taken that the participants would not see each other. All participants in the identifiable condition agreed by signing the consent form to have their photo taken with a Skype snapshot, so that the photo would appear along side with their name, which was entered earlier by the experimenter. The study was all computer-based and participants completed at first a pre-questionnaire online [27]. At its end, they were asked to click on a Skype tab at the bottom of the screen.

This synchronous communication setting was used to introduce the two aspects of online interactivity, namely users being able to share content and to receive feedback on it from others. Participants were told that they would interact with a group that advocated the protection of the environment on campus, so a movement organization. It was further mentioned that all other users agreed that it was important to protect the environment, just like they did themselves. This provided information about the group norms and seemed plausible as it appeared after the participants answered questions about their opinion on the need to protect the environment. Analyses of these ratings showed that participants' initial (t1) mean attitude score ($M = 5.15$ ($SD = .72$)) was significantly above the median of the scale (m = 4) (t (83) = 14.58, $p < .05$), indicating that the feedback should have not seemed contradictory.

A confederate made sure that participants only started sharing content once all participants were active in the Skype chat and she served as a representative of the movement organization, repeating its goal to protect the environment. Participants

further contributed to each other's knowledge about the achievements made to act more sustainable on campus (see below). The confederate prompted the general topics to which ideas or reports of behavior could be shared by participants. Those were their thoughts about how the environment is treated, what could be done to improve the state of the environment, what was already done on campus (performance information) and how more people could get involved in these activities. Participants were encouraged to refer to and comment on each others' contributions, raising the salience of an active reward system. After 15 minutes, participants filled in a second questionnaire and then went back to the Skype chat for another 10 minutes before filling in the post-questionnaire. Participants were then thanked and debriefed.

4.4 Measures

The following scales were administered in the study. The willingness to participate in offline collective actions was assessed with four self-designed items (e.g., "I am open to engage in this demonstration.", "I would initiate even more such panel discussions."), after describing two different collective action scenarios: a demonstration and a panel discussion. In each case, members of an organization that had goals different from those of the participants' organization (an outgroup) were described to be present. The goal of the panel discussion was the consensus of both organizations. For the demonstration the necessity of a compromise of the group's aims was not mentioned. On the contrary, conflicts were predicted to arise. This distinction was chosen to explore whether the effects of empowerment varied for different types of collective actions. Reliabilities of both scales were acceptable at all three measuring points (demonstration: t1 α =.88, t2 α =.88, t3 α =.86; panel discussion: t1 α = .82, t2 α = .82, t3 α = .84).

Psychological empowerment for engaging in collective actions was measured with Spreitzer's [15] scale that includes 12 items (e.g., "I am confident about my ability to take part in this action.", "My impact on what happens during this action is large."), corresponding to the concepts of meaning, competence, self-determination and impact. As the original items were phrased for a work context, we adapted the wording to fit the setting of collective actions. Empowerment was assessed for both scenarios of collective actions. The reliabilities of the scale were acceptable (panel discussion: t1 α =.88, t2 α =.87; demonstration: t1 α =.91, t2 =.88).

A number of control variables were measured as well. Participants' attitude towards the need to protect the environment was analyzed with the New Environmental Paradigm Scale (14 items) [25]. It assesses perspectives on environmental protection and industrial growth (e.g., "The balance of nature can be easily upset by the actions of humans.", "Humans do not need to adapt to the natural environment because they can remake it to suit their needs." (reversed)). Reliabilities at all three measuring points were acceptable (t1 α =.77, t2 α =.76, t3 α =.83). Further, identification with the movement organization was measured. In previous research, identification with the group on which's behalf one would participate in a collective action has been a robust predictor of participation [1]. More precisely, identification with the movement itself and not the social category it represents has predicted engagement [26]. Identification was operationalized by the mean agreement with Krizan and Baron's five-items scale [27] (e.g., "I respect the members of my group.", "Overall, I identify with

the members of my group.") (reliabilities: t1 α =.75, t2 α =.93). We also administered a scale for perceived social support from the movement organization with an adapted five-item version of the multidimensional scale of perceived social support [28] (e.g., "The group members could be a source of support for me.", "The members of my group would be willing to help me make decisions.") (reliability: t2 α =.91). Further, perceived obligation to act according to what the group values and perceived identifiability of the other participants were measured with one self-designed item each.

4.5 Results

A manipulation check showed that participants perceived others indeed to be more identifiable in the condition where a name and photo of the participants was provided ($F(1, 82) = 81.57$, $p < .05$). Descriptive information of all dependent and independent variables at all three measuring points is available in Table 1.

Table 1. Descriptive information

Variable	Time	*M*	*SD*
Participation in a demonstration	t1	3.98	1.36
	t2	4.19	1.38
	t3	4.31	1.47
Participation in a panel discussion	t1	4.63	1.12
	t2	4.57	1.09
	t3	4.74	1.25
Empowerment (demonstration)	t1	3.71	1.19
	t2	4.21	1.13
Empowerment (panel discussion)	t1	4.08	1.07
	t2	4.45	1.03
Obligation	t2	4.39	1.59
Social support	t2	4.81	1.14
Identification	t1	5.12	.72
	t2	6.01	.98
Attitude	t1	5.14	.72
	t2	5.56	.71
	t3	5.63	.80

A repeated measure analysis, using the scores of willingness to participate in a demonstration as a within-subject factor with three levels (t1 to t3), indicated a significant within-subject effect ($F(1, 83) = 7.93$, $p < .05$). For the panel discussion, no significant within-subject effect was shown ($F(1, 83) = .75$, $p > .05$).

A multivariate analysis of variance, using the experimental conditions as a between-subject factor, showed a significant difference between the conditions for both types of collective actions at t3, with willingness to engage in collective actions being higher when participants were identifiable (demonstrations: $F(1, 82) = 3.76$, $p = .05$; panel discussion: $F(1, 82) = 4.22$, $p < .05$). Table 2 shows further the respective between subject differences for empowerment, perceived social support, attitude and perceived obligation, with scores being always significantly higher in the identifiable condition.

Table 2. Between subject differences on independent and control variables

Variable	time	df (between-, within-group)	*F*	*p*
Empowerment (demonstration)	t1	1, 82	6.25	< .01
	t2	1, 82	4.73	< .05
Empowerment (panel discussion)	t1		5.28	< .05
	t2	1, 82	7.89	< .01
Social support	t2	1, 82	7.93	< .01
Identification	t1	1, 82	3.13	.08
	t2	1, 82	19.07	< .01
Obligation	t3	1, 82	10.63	< .01
Attitude	t1	1, 82	.99	.32
	t2	1, 82	3.25	.08
	t3	1, 82	2.56	1.13

A repeated measure analysis, using the t1 and t2 scores of identification with the group as within-subject factors, showed further that identification increased significantly over time ($F(1, 82) = 66.12$, $p < .05$). Attitudes on the need to protect the environment also increased significantly from t1 to t2 ($F(1, 82) = 82.94$, $p < .05$).

Hypothesis 1. Empowerment to participate in a panel discussion increased significantly between t1 and t2 ($F(1, 82) = 12.74$, $p < .05$). Empowerment to participate in a demonstration also increased significantly from t1 to t2 ($F(1, 82) = 21.90$, $p < .05$). Hypothesis 1 was not rejected.

Hypothesis 2. Across conditions, two multiple step-wise regressions, with the mean scores of willingness to participate in a panel discussion or demonstration (t3) as dependent variables and the respective empowerment score (t2) as independent variable, controlling for identification (t2), attitude (t2), social support (t2), felt obligation (t2) and initial participation ratings (t1, t2) showed that the willingness to participate in panel discussions was predicted by higher empowerment ($F(2, 81) = 41.23$, $\beta = .70$, $p < .05$, $R^2 = .51$) and a less extreme attitude towards the need to protect the environment ($F(2, 81) = 41.23$, $\beta = -.19$, $p < .05$, $R^2 = .51$). Higher willingness to participate in demonstrations was predicted by higher empowerment ($F(1, 82) = 103.44$, $\beta = .75$, $p < .05$, $R^2 = .56$). Hypothesis 2 was not rejected.

Hypothesis 3. Given those findings, we could test the effect of the experimental condition on the willingness to participate in both types of collective action (t3) mediated by the sense of empowerment (t2). Bootstrap analysis using 10 000 bootstraps showed for the panel discussion (LL (95): .14, UL (95): .88, indirect effect: .49) and demonstration (LL (95): .06, UL (95): .98, indirect effect: .51), that being identifiable when contributing content and receiving feedback on it led to higher willingness to participate in collective action, through a higher sense of empowerment. Hypothesis 3 was not rejected.

5 Discussion

The goal of this article was to revisit current research on the impact of Internet use on participation in collective actions, by considering the element of online interactivity that defines the participatory Internet. We showed that giving users the opportunity to share their ideas and letting them receive feedback on them, increased their sense of empowerment to carry out a collective action on behalf of an environmentalist movement organization, which led to higher willingness to participate for the movement organization in a panel discussion and demonstration. This finding is in line with earlier work and supports the positive effect of Internet use on engagement in civic and political actions offline [2]. The influence was shown for actions that ask the user to compromise the goals of the organization she or he represents and such where potential conflicts with opposing groups might arise.

The results challenge the frequently deployed instrumental approach [5]. This focuses solely on the costs of information access, while our findings turn the attention towards intra-individual processes that Internet use can affect. We showed that one's sense of meaning, competence, self determination and impact [15] can be increased through Internet use, and can in return drive behavioral intentions.

The effect of empowerment on the willingness to participate in collective actions depended further on the communication context. Being identifiable with one's name and photo to other users, increased one's sense of empowerment and thereby the behavioral intention to carry out collective actions offline. This emphasizes the role of the reward and incentive system as an antecedent of one's sense of empowerment [15]. It seems to be more effective if the user's contributions are personally related to

her or him. Then, the user's motivation for a positive self-presentation should be heightened and result in more rewarding feedback, as more opinions and behavior in line with the group's norms should be expressed [22]. Indeed, the perceived obligation to behave in the way that is appreciated by the movement organization was higher in the identifiable condition.

One might argue that higher perceived obligation would not just affect the kind of content shared online, but also the participants' ratings for empowerment and willingness to participate in collective actions, which can be seen as a behavior in line with the group norms. However, these measures were administered in an anonymous questionnaire and it was at no point mentioned that they would be transmitted to the group. In addition, obligation did not predict the dependent variables significantly.

Further alternative explanations of our findings could be based on another mechanism that the experimental conditions might have evoked. In the interaction settings, users were not only identifiable to the others they shared content with, but those others were also identifiable to them. Previous research indicated that identifiability of an audience online can increase the users feeling of social support and solidarity in the group [22]. Indeed, perceived social support was significantly higher in the identifiable condition, but it was not a significant predictor of willingness to participate in collective actions once empowerment was controlled for.

While earlier work often argued that Internet use can mobilize especially users that are already motivated to engage in actions, we showed that intrinsic task motivation [16] can actually be increased through Internet use itself, in particular through social media that allows users to generate content and receive feedback on it. This widens the influence of the Internet to an audience with different levels of initial willingness to engage in an offline action on behalf of a movement organization.

Design principles can be derived based on our findings, enabling movement organizations to empower its users and to increase their willingness to get involved for the organization. The elements that were shown to affect the users' willingness to engage in collective actions offline are such that enable users to generate content and receive feedback on it while being identifiable with personal attributes such as one's name or photo. An example for this can be found on the website of Greenpeace International (*www.greenpeace.org*). Real time reports document when online petitions were signed or donations were made by users, which are identifiable with their name.

Future studies could extend our findings by assessing the differential effects of identifiability, namely that of users and the audience, so that a clear distinction of the effects of empowerment and social support can be made. In addition, other intra-individual processes that can be influenced by Internet use and in return affect collective action participation should be investigated; for example the phenomenon of group polarization. It refers to attitudes becoming more extreme, in direction of the initial opinion [29] as the result of an interaction with like-minded others and based on processes of social comparison or persuasive arguments [30]. Previous work indicated that more extreme attitudes predicted behavior more strongly than moderate ones and particularly such behavior that is less compromising [31]. Also in our study attitudes shifted from an already supportive stand point to one that expressed a more extreme need to protect the environment. This more extreme position predicted a lower

willingness to participate in a collective action that had as a goal a compromise with the outgroup. Furthermore, a more realistic setting, for example in the context of an existing movement organization, would enhance the ecological validity of the findings. While we were able to support our predictions using a synchronous communication mode, it would be interesting to assess them in an asynchronous one as well, as this defines many interactive features online, such as comments on status updates or blog posts. Having only female participants restricts the generalization of our results and the gender distribution should be more balanced in upcoming research. Despite these limitations, our study opens the field for new research questions that focus strongly on the user and effects on intra individual processes due to the fact that social media allows users to be actively involved in generating online content.

References

1. Postmes, T., Brunsting, S.: Collective Action in the Age of the Internet: Mass Communication and Online Mobilization. Social Science Computer Review 3, 290–301 (2002)
2. Boulianne, S.: Does Internet Use Affect Engagement? A Meta-Analysis Research. Political Communication 26, 193–211 (2009)
3. Wright, S.C., Taylor, D.M., Moghaddam, F.M.: Responding to membership in a disadvantaged group: From acceptance to collective protest. Journal of Personality and Social Psychology 58, 994–1003 (1990)
4. Rafaeli, S., Ariel, Y.: Assessing Interactivity in Computer-Mediated Research. In: Joinson, A., McKenna, K., Postmes, T., Reips, U.-D. (eds.) The Oxford handbook of Internet Psychology, pp. 165–186. Oxford University Press, New York (2007)
5. Xenos, M., Moy, P.: Direct and Differential Effects of the Internet on Political and Civic Engagement. Journal of Communication 57, 704–718 (2007)
6. Van Laer, J., Van Aelst, P.: Cyber-protest and civil society: the Internet and action repertoires in social movements. In: Jewkes, Y., Yar, M. (eds.) Handbook of Internet Crime, pp. 230–254. Willian Publishing, Devon (2010)
7. Wall, M.: Social Movements and email: expressions of online identity in the globalization protests. New Media & Society 9, 258–277 (2007)
8. Norris, P., Jones, P.: Virtual Democracy. Harvard International Journal for Press/Politics 293, 474–480 (1998)
9. Bimber, B.: The Internet and Citizen Communication with Government. Does the Medium matter? Political Communication 16, 409–429 (1999)
10. Krueger, B.S.: Assessing the potential for Internet political participation in the United States: A resource approach. American Politics Research 30, 476–498 (2002)
11. Fisher, D.R., Boekkooi, M.: Mobilizing friends and strangers: Understanding the role of the internet in the Step IT UP day. Information, Communication & Society 2, 193–208 (2010)
12. Gil de Zúniga, H., Valenzuela, S.: The Mediating Path to a Stronger Citizenship: Online and Offline Networks, Weak Ties, and Civic Engagement. Communication Research 10, 1–25 (2010)
13. Granovetter, M.S.: The strength of weak ties. American Journal of Sociology 78, 1360–1380 (1973)

14. Kavada, A.: Web 2.0 platforms and Political Communication: Conceptual and methodological challenges. In: Annual meeting of the International Association for Media and Communication Research (2010)
15. Spreitzer, G.: Psychological Empowerment in the Workplace: Dimensions, Measurement, and Validation. Academy of Management Journal 5, 1442–1465 (1995)
16. Thomas, K.W., Velthouse, B.A.: Cognitive elements of empowerment. Academy of Management Review 15, 666–681 (1990)
17. Wallace, J.C., Johnson, P.D., Mathe, K., Paul, J.: Structural and Psychological empowerment Climates, Performance, and the Moderating Role of Shared Felt Accountability: A Managerial Perspective. Journal of Applied Psychology 4, 840–850 (2011)
18. Seibert, S.E., Wang, G., Courtright, S.H.: Antecedents and Consequences of Psychological and Team Empowerment in Organizations: A Meta-Analytic Review. Journal of Applied Psychology 5, 981–1003 (2011)
19. Kantner, R.M.: The new Managerial work. Harvard Business Review 66, 85–92 (1989)
20. Reicher, S.D.: Collective action and the (re)construction of the self. In: Oosterwegel, A., Wicklund, R. (eds.) The self in European and North American Cultures: Development and Processes. Kluwer, Amsterdam (1994)
21. Turner, J.C., Hogg, M.A., Oakes, P.J., Reicher, S.D., Wetherell, M.S.: Rediscovering the social group. A self-categorization theory. Blackwell, Oxford (1987)
22. Spears, R., Lea, M.: Panacea or Panopticon?: The hidden Power in Computer-Mediated Communication. Communication Research 21, 427–459 (1994)
23. Postmes, T., Spears, R., Lea, M.: Breaching or building social boundaries? SIDE-effects of computer-mediated communication. Communication Research 25, 689–715 (1998)
24. Leiner, D.J.: oFb – the online questionnaire (2011), `https://www.soscisurvey.de`
25. Dunlap, R.E., Van Liere, K.D.: The new environmental paradigm. Journal of Environmental Education 9, 10–19 (1978)
26. Simon, B., Loewy, M., Stürmer, S., Weber, U., Freytag, P., Habig, C., Kampmeier, C., Spahlinger, P.: Collective identification and social movement participation. Journal of Personality and Social Psychology 74, 646–658 (1998)
27. Krizan, Z., Baron, R.S.: Group polarization and choice-dilemmas: How important is self-categorization? European Journal of Social Psychology 37, 191–201 (2007)
28. Zimet, G., Dahlem, N., Zimet, S.G., Farley, G.: The Multidimensional Scale of Perceived Social Support. Journal of Personality Assessment 52, 30–41 (1988)
29. Lamm, H., Myers, D.B.: Group-induced polarization of attitudes and behavior. In: Berkowitz, L. (ed.) Advances in Experimental Social Psychology, vol. 2, pp. 147–195. Academic Press, New York (1978)
30. Burnstein, E., Vinokur, A.: Persuasive argumentation and social comparison as determinants of attitude polarization. Journal of Experimental Social Psychology 13, 315–332 (1977)
31. Petty, R.E., Krosnik, J.A.: Attitude strength: antecedents and consequences. Lawrence Erlbaum Associates, Mahwah (1995)

The Illusion of Agency: The Influence of the Agency of an Artificial Agent on Its Persuasive Power

Cees Midden and Jaap Ham

Human-Technology Interaction, Eindhoven University of Technology,
P.O. Box 513, 5600 MB Eindhoven, The Netherlands
c.j.h.midden@tue.nl

Abstract. Artificial social agents can influence people. However, artificial social agents are not real humans, and people may ascribe less agency to them. Would the persuasive power of a social robot diminish when people ascribe only little agency to it? To investigate this question, we performed an experiment in which participants performed tasks on a washing machine and received feedback from a robot about their energy consumption (e.g., "Your energy consumption is too high"), or factual, non-social feedback. This robot was introduced to participants as (a) an avatar (that was controlled a human in all its feedback actions; high agency), or as (b) an autonomous robot (that controlled its own feedback actions; moderate agency), or as (c) a robot that produced only random feedback; low agency). Results indicated that participants consumed less energy when a robotic social agent gave them feedback than when they received non-social feedback. This behavioral effect was independent of the level of robotic agency. In contrast, a perceived agency measure indicated that the random feedback robot was ascribed the lowest agency rating. These results suggest that the persuasive power of robot behavior is independent of the extent to which the persuadee explicitly ascribes agency to the agent.

Keywords: Persuasive Technology, Agency, Social Robotics, Persuasive Power.

1 Introduction

Imagine the near future in which one of your two personal service robots compliments you on how great you look today. A few minutes later, your other robot approaches you. This second robot has been malfunctioning since the weekend, uttering random sentences. In other words, it clearly is no longer an independent agent that controls what is says. The latter robot has lost some if not most of its agency, that is, it seems to say things involuntarily. You keep reminding yourself that you should hook it up to its maintenance website. While passing, this second robot gives you a compliment on your pleasant demeanor this morning. Now, which of the two robots would receive your warm appreciation in return to its compliment? Would your knowledge about the lowered agency of robot number two decrease the persuasive power of that robot?

M. Bang and E.L. Ragnemalm (Eds.): PERSUASIVE 2012, LNCS 7284, pp. 90–99, 2012.

Earlier research indeed indicates that artificial social agents can have persuasive power. For example, research indicates that when a robotic agent gives participants feedback on their energy consumption in a task in which participants have to make energy consumption decisions, its feedback influences the amount of energy participants consume [5]. More specifically, social feedback from a robotic social agent had stronger persuasive power than comparable feedback coming from a non-social source.

However, artificial social agents are not real humans, and people may ascribe different amounts of agency to them, which can be described as the extent to which a human has the feeling to have interaction with a sentient human being [2, 4]. Although it seems plausible that perceived agency is important to make social behavior of an agent, whether human or non-human, meaningful, recent research has produced some inconsistent results and different views on the effects of agency [see e.g. 9]. In the present paper we will present new research about the effect of agency on the persuasive power of a social agent. First we will briefly review the current state of knowledge regarding social responses to artificial social agents.

Earlier theorizing about the role of embodied virtual agents (EVA's) in social interaction has led to two views (relevant for robot agency effects) that stand out and have been scrutinized empirically in two lines of research. The first view has resulted from the work by Reeves and Nass [15]. Their research suggests that people exhibit social responses when interacting with a computer [e.g., 3, 6, 7, 8, 12, 13]. According to the media equation theory [15], people consider computers as social actors when triggered by certain social cues. These responses are described in the Computer As Social Agents (CASA) paradigm [15]. This paradigm proposes that people react to computers as if they are social actors, while at the same time, people know that these responses are inappropriate (i.e., people know they do not have to react socially to a computer). Numerous studies have supported this notion [for an overview see 15]. For example, it has been shown that people react more politely when a computer asked about its own performance, than when the computer's performance was assessed by a paper-and-pencil questionnaire [13], which is similar to reactions in human-human interaction. However, when asked about their social reactions to the computer, people often deny behaving socially and may even feel offended by the question. In other words, people's implicit reactions (behavior when interacting with computers) and explicit responses when directly asked about computers differ from each other.

Based on such earlier findings, Reeves and Nass hypothesized that humans react in an *automatic* fashion to technical systems if social cues are provided by the system [15]. This explanation is based on a theory by Langer [10] about mindlessness behavior. When in a state of mindlessness, people act on automatic pilot and do not think consciously about their behavior. Attention is directed to information that is relevant for the current task, leaving no room for alternative information [10]. It is suggested that when people interact with computers, they mindlessly form a premature commitment to social scripts of human-human interaction experienced in the past [12]. In short, the CASA paradigm [15] suggests that social cues trigger social rules and people react in accordance with these social rules. However, only few studies tested the effects of social human-agent interactions on persuasion and

behavior change. Taken together, the CASA paradigm has demonstrated important psychological effects of human-computer interactions that include social aspects similar to interactions between humans. The hypothesis of automaticity, or mindlessness as suggested by Nass and colleagues [e.g., 12] thereby referring to Langer [10] deserves to be put under empirical scrutiny.

Research investigating the CASA paradigm usually used simple boxed computers, thereby demonstrating the easiness of evoking social responses. This raised the question of how features of the agent that make the agent more realistic, that is, more human-like and more social, moderate the amount of social responses that are triggered in humans interacting with those more human-like artificial agents. This has been the focus of the second type of view on the effects of agency.

Different from the CASA paradigm, this other view has a focus on embodied virtual agents, which usually have humanlike features. For example, the Threshold Model of Social Influence [2] proposes that social human-agent interaction is a function of perceived agency and the social realism of the agent. This model assumes that people react socially to humans or avatars on the basis of social verification, that is, the feeling that they are engaging in a meaningful, social interaction. High agency (perceived for example when interacting with an avatar) makes social realism less relevant. When agency is low, however, social realism may be needed for a human actor to cross the threshold and engage in a social way with an artificial agent. Social realism may be dependent on the behavior and the appearance of the agent [2].

In support of the Threshold Model of Social Influence, earlier research has demonstrated some effects of agency [1] and of social realism [14, 9]. But these results are not consistently found. For example, Vossen, Midden and Ham [17] showed that both speech and appearance of a persuasive agent served as social cues and promoted social influence. However, the two cues did not add up. One cues appeared sufficient to create the effect, suggesting that the cues worked as triggers to start up a social script. Neither is there evidence for the supposed interaction between social realism and agency [see for an overview 9]. Both Nass and Moon [12] and Kraemer [9] indicate that the behavior of the agent is more important than its social features.

In sum, two contrasting conclusions about the relationship between (perceived) agency and the extent to which people engage in a social way in the interaction with an artificial agent can be drawn from these theories. According to the CASA paradigm it does not matter how much agency an agent has, because people will automatically interact with the agent in a social way as long as it shows social cues. However, according to the Threshold Model [2] the effect of an artificial agent will be dependent on its perceived agency and, if perceived agency is not high enough, also on the social realism of the agent. Based on more recent work both models suggest that the richness of social behavioral cues makes a difference on people's social responses. However, differences in research paradigms and design limitations make it difficult to draw clear conclusions about the mechanisms underlying social responses in human-agent interactions. Furthermore, most effects concern psychological effects in terms of perceived presence, liking of the agent and so on. Only very few studies have addressed the potential of artificial agents to influence human behavior through social influence.

1.1 The Current Research

Our experiment was primarily designed to test the effects of agency on behavior change through social feedback. Secondly, we want to compare the effectiveness of feedback by a social realistic agent to the effectiveness of feedback by a non-social system. Thirdly, we wanted to test whether behavioral changes were mediated (or not) by conscious judgments about the attributed social features of the agent .

The research setup was based on earlier research by Midden and Ham [11] in which a robotic agent provided social feedback. In various studies it was demonstrated that social feedback (through social approval and disapproval) was effective in enhancing the saving of energy in household tasks like programming a washing machine. Moreover, by comparing the social feedback to non-social evaluative feedback the behavioral effects could be attributed to the perceived socialness of the agent. Also in the present research a robotic agent (Philips iCat) was used to give social feedback on energy consumption, thereby trying to convince people to use less energy whereas the participants also have the additional goal of producing a clean laundry. The current experiment had four conditions. In the first condition, the robot had the highest level of agency, that is, the robotic agent was functioning as an avatar that was fully controlled by a human actor. In the second condition, the robot had somewhat lower level of agency, that is, the same robotic agent was seemingly acting autonomously. In the third condition, the robot has the lowest level of agency, that is, the robotic agent was controlled (as presented to the user) by a lottery system. Finally, there was a control condition in which the feedback was offered through an energy meter that indicated energy consumption on a simple bar that changed height dependent on energy consumption levels [see also 11]. This latter feedback was also evaluative because the bar could for example reach the top of the scale, or only the middle of the scale, just as the feedback given in all three robot conditions, but was inherently not social in nature. Thereby, social cues were completely absent in the feedback source (the energy meter) used in this condition.

Our research assessed two dependent variables. Firstly, we measured the user's (virtual) energy consumption, which functioned as a measure of persuasion by the feedback source (one of the three types of robot or the energy meter). Secondly, we measured the explicit judgments participants made about their perceived agency and social characteristics of the robotic agent.

This design made it possible to compare different levels of agency, to compare the 'social' conditions with the 'non-social' control group and to test whether the behavioral effects were mediated by conscious judgments or not.

2 Method

2.1 Participants and Design

Eighty-four participants (55 men and 29 women, average age = 21.4 years, SD = 3,5) were randomly assigned to one of four experimental conditions: high agency condition, moderate agency condition, low agency condition, or the non-social

feedback condition. The experiments lasted 30 minutes for which participants were paid five Euros. All participants were students at Eindhoven University of Technology.

2.2 Procedure and Materials

Participants were invited to engage in an experiment using a simulated washing machine, and were seated individually in a small room. For participants in robotic feedback conditions, an iCat was positioned on the participants' desk, next a computer. An iCat is a robot developed by the Philips Corporation in the form of the stylized head of a cat that is able, among others, to display social expressions by moving lips, eyes, eyelashes, and eyebrows, and by playing speech files. For participants in the non-social feedback condition, the iCat was not present. For all participants, a simulated washing machine panel was presented in the top half of the screen (see Figure 1). Only for participants in the non-social feedback condition, we added an energy meter to the panel. This energy meter clearly displayed the amount of electricity in kWh corresponding to the chosen washing program. For all participants, in the bottom half of the screen, a program displaying the instructions, tasks and questions was presented.

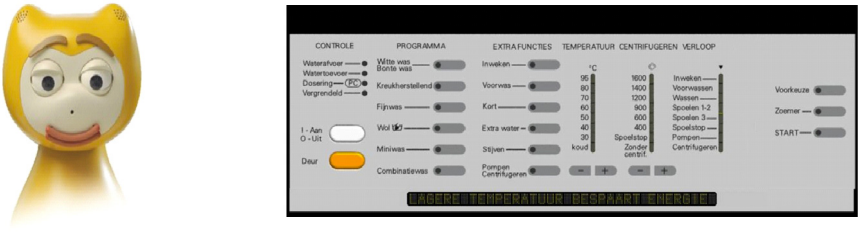


Fig. 1. The iCat and the simulated washing machine interface

This program started with general introductions, and then instructed participants about the task: they were asked to complete several simulated washing trials on a computer. Next, participants were instructed on how to program the washing machine. Participants then were instructed to do each washing as good as possible, that is, to clean the clothes and not damage them (e.g., by washing a woolen sweater at 90°C, causing it to shrink), but also to use as little electricity as possible. So, participants were given two (partly contrasting) goals. To save energy, other settings are optimal (e.g., washing cold) than to clean clothes. Thereby, we motivated participants to use (some) energy, while installing also the motive to save energy for which we could then provide feedback.

To participants in the non-social feedback condition the program next explained the energy meter. To participants in the robotic feedback conditions, the program explained that the iCat would help them save energy by giving feedback on their energy consumption. In high agency condition, a man introduced himself in a video message presented on screen, explaining he would operate the iCat. In the moderate

agency condition, the program explained that the iCat was an independent robot, and would give its own assessments in its feedback on the energy consumption of the participant. In the low agency condition, the program explained that chance would determine which feedback would be given, and this was illustrated with an example.

Next, all participants performed a series of washing trials. After one practice trial, the actual 10 washing trials started. For each trial (and also the practice trial), participants were instructed to complete a specific type of wash (e.g., "wash four very dirty jeans"). Each description of a specific type of wash was randomly drawn from a collection of thirty descriptions of common washes, for each trial of each participant such that each participant completed ten different washes. During each washing trial, participants were able to change settings on the washing machine panel until they were satisfied and then pressed a "start" button.

Participants received feedback about the energy consumption of the chosen washing program after each change of settings. Feedback depended on the difference between the chosen settings of the washing machine and a predefined value for that specific washing task. The size of this difference determined the feedback level. We introduced some noise to this difference (by adding a random factor) to the feedback to make the relation between a participant's settings and feedback about their energy consumption less transparent. Thereby we were able to provide also participants in the random robotic feedback condition with correct feedback (at least part of the time) without enhancing perceived agency. In the introduction of the experiment the participants were told how their feedback system worked. For example, they were told the feedback was generated randomly. However, the actual feedback given was identical in all three conditions. Thus, in our setup we kept constant the behaviors of the agents in the three agency conditions.

Participants in the non-social feedback condition received this feedback through the energy meter. The energy meter gave feedback by indicating energy consumption on a scale (on the positive side labeled with "high", "normal", and "low", and on the negative side labeled with "low", "normal", and "high"). Likewise, participants in three social robotic feedback conditions received feedback through the iCat during each trial. For small differences from average energy consumption, the iCat uttered a moderately positive or negative word ("Alright", or "Pretty bad"), for larger deviations a more gravely positive or negative word ("Very good", or "Very bad"), and for large deviations the iCat uttered an even more gravely positive or negative word ("Fantastic", or "Terrible"). All presented labels are best possible translations from Dutch.

After all 10 washing trials had been completed, we assessed the agency judgments of the participants in the robotic feedback conditions ascribed to the iCat. For this, these participants evaluated the iCat on seventeen bipolar 7-point semantic differentials questions. For each of these questions, a word related to low agency was presented on the left of the scale, and a word related to high agency was presented on the right of the scale. Combinations were for example "stupid" and "smart," "non-expert" and "expert," "not social" and "social," and "no intentions" and "full of intentions". A factor analysis of these questions indicated three components with eigenvalues above 1. Only the second (12% explained variance) of these components

was clearly related to agency (as it loaded on questions relating to initiative and intentionality). The following items had factor loadings >. 50: passive-active, non-social-social, unconvincing-convincing, untrustworthy-trustworthy, unsupportive-supportive, dead-alive, We used this factor as the agency score in our analyses. The first component (30.2% explained variance) was related quite clearly to questions that tapped perceived competence. The third component (6,6% explained variance) was not clearly related to specific questions. Finally, participants answered demographic questions, were debriefed and thanked for their participation.

3 Results

For each of the ten washing trials of each participant, we calculated the difference between the amount of electricity a participant's settings would have used, and the average usage of electricity for that specific type of wash (e.g., the 4 very dirty jeans) by all participants in our study. We labeled this the energy consumption score. This way, we were able to calculate a dependent variable that indicated the difference between a reference amount of electricity needed for a specific type of washing task (at least in the current study) and the electricity a participant chose to use. To be able to distinguish the effects of positive and negative feedback we calculated an index based on total number of actions of users in the user interface. This means that we not only included the final choices per trial, but all the preceding programming choices. As explained in the method section, these were all followed by (non-social or social) feedback, either positive or negative. The index subtracted for each action the following choice, in terms of energy consumption effect, from the current choice, thereby indicating whether the feedback resulted in a higher or lower energy consumption score for the next following choice.

These scores were submitted to a 4 (feedback condition: high agency feedback vs. moderate agency feedback vs. low agency feedback vs. non-social feedback) x 2 (feedback valence: positive vs. negative) mixed model analysis. This analysis indicated the expected main effect of type of feedback, $F(1, 76.343) = 4.31$, $p < .01$. More specifically, this analysis indicated that non-social feedback resulted in a higher energy consumption score ($M = .09$, $SD = .02$) than in all of the robot feedback conditions, all p's < .03, and presented no evidence that feedback by a robot guided by a human agent (high agency condition) resulted in a different energy consumption score ($M = .04$, $SD = .02$) compared to feedback by an independent robot (moderate agency condition; $M = .01$, $SD = .02$), or feedback by a robot that provided random feedback (low agency condition; $M = .02$, $SD = .02$), all p's > .20.

In line with earlier research suggesting the stronger persuasive effects of negative feedback (a.o. Midden & Ham, 2009), this analysis also indicated that negative feedback resulted in a lower energy consumption score ($M = -.08$, $SD = .01$) than positive feedback ($M = .16$, $SD = .01$), $F(1, 3659.411) = 381.28$, $p < .0001$. This analysis did not indicate an interaction of feedback valence x feedback condition, $F(1, 3654.520) = 1.19$, $p = .31$, indicating that the stronger effect of negative versus positive feedback did not differ for the four types of feedback that we gave participants (non-social, and high, moderate, or low agency feedback).

To analyze effects of the type of agent on participant's agency judgments, we compared participant's agency scores in 1 x 3 (feedback condition: high agency feedback vs. moderate agency feedback vs. low agency feedback) ANOVA. This analysis indicated a main effect of feedback condition, $F(2, 61) = 3.17$, $p < .05$. More specifically, specific contrast analysis indicated that participants' agency judgment scores in the high agency condition ($M = -.28$, $SD = .89$) and in the moderate agency condition were higher ($M = -.27$, $SD = .79$) than participants' agency judgment scores in the low agency condition ($M = -.79$, $SD = .67$), $F(1, 61) = 6.33$, $p = .02$. No evidence was found that participants' agency judgment scores in the high agency condition and in the moderate agency condition differed, $F < 1$. Furthermore, both participants' agency judgment scores for the avatar feedback system differed from participants' agency judgment scores for the random agent feedback, $F(1, 61) = 4.77$, $p < .05$, as did their judgment score for the robot feedback system, $F(1, 61) = 4.66$, $p < .05$. Finally, results indicated no correlation between participants' agency judgment scores and their energy consumption scores, $p > .43$., indicating that agency judgments did not mediate the effects of our agents manipulation on energy consumption behaviors.

4 Conclusion and Discussion

The aim of this research was to investigate whether the persuasive power of an artificial social agent would diminish when the persuadee ascribed less agency to it. To investigate this question, we conducted an experiment in which participants performed tasks on a washing machine and received feedback from a robotic artificial agent about their energy consumption (e.g., "Your energy consumption is too high"), or,non-social feedback. This robot was introduced to participants as (a; high agency condition) an avatar (a human completely controlled all its feedback actions), or as (b; moderate agency condition) an autonomous agent (that controlled all its own feedback actions), or as (c; low agency condition) an agent that (in the eye of the user) provided only random feedback. Results indicated that participants consumed different amounts of energy dependent on the type of feedback. Inspection of the specific contrasts revealed that the greater savings were achieved in the three conditons with social feedback compared to the non-social feedback condition. However no differences in behavioral effects were found between the three agency levels.

Furthermore, results indicated that participants ascribed less agency to the random feedback agent than to the other two agents. However, the difference in perceived agency between avatar and autonomous agent did not reach significance. This finding confirms that, at least at a conscious level, participants were aware that the random feedback agent had less agency than the other two robots. Notwithstanding this awareness of diminished agency, results did not provide evidence that the influence of the social feedback provided by each of the three robots on participants energy consumption choices differed. Furthermore, we could not establish a mediation effect of (explicitly) perceived agency on behavior.

Thereby, the current results provide us with more insight in the cognitive processes of persuasion by artificial agents. Our results indicate that the effect of our three-levels agency manipulation on the behavioral impact of social feedback by the agent did not require an explicit awareness of the level of agency. Furthermore, the agency manipulation did not show a difference at the behavioral level. Apparently, the experience of the agent providing spoken social feedback, as contrasted to the non-social feedback condition, was sufficient to produce the observed behavioral effects. This result seems to support the automaticity hypothesis as proposed in the CASA paradigm [15].

What can be said about the role of social realism as, among others, suggested by the Threshold Model of Social Influence [2]? Our design did not allow for directly testing interaction effects of agency and social realism, but the comparison of the three social agent conditions with the non-social control condition suggests that the higher social realism of the agent's behaviors and appearances compared to the lack of social realism in the control condition (the energy meter gave non-social feedback) contributed to the behavioral effects of the feedback.

In sum, the current research suggests that explicit knowledge indicating higher or lower perceived agency of an artificial agent does not lead to an increase or a decrease of the persuasive power of that agent. The socialness of the agent seems to be more important in cueing the user to activate a social interaction mode. Future research could investigate the additional effects of determinants of social realism of robotic agents. In addition further insight is needed into the different processing modes of assessing agency and social realism by the human user. To return to the description of the two personal service robots you might find present in your home in the near future. The current research suggests that both robots might receive your warm appreciation when they utter a compliment, even when you know that the second robot desperately needs maintenance because it only utters random sentences. Even when an agent seems to have little agency, the social cues that it uses to persuade could be effective, and be able to influence persuadees through automatic, unconscious cognitive processes.

Acknowledgments. We would like to thank Corine Horsch and Bo Merkus for executing the research, and the TUe-HTI Persuasive Technology Lab group for scientific reflection. This research was sponsored by the Dutch scientific research grants organization NL Agency.

References

1. Aharoni, E., Fridlund, A.J.: Social reactions toward people vs. computers: How mere labels shape interactions. Computers in Human Behavior 23, 2175–2189 (2007)
2. Blascovich, J.: A theoretical model of social influence for increasing the utility of collaborative virtual environments. In: Proceedings of the 4th International Conference on Collaborative Virtual Environments, pp. 25–30. Association for Computing Machinery, Bonn (2002)

3. Fogg, B.J., Nass, C.: Silicon sycophants: the effects of computers that flatter. International Journal of Human-Computer Studies 46, 551–561 (1997)
4. Guadagno, R.E., Blascovich, J., Bailenson, J.N., McCall, C.: Virtual humans and persuasion: The Effect of agency and behavioral realism. Media Psychology 10, 1–22 (2007)
5. Ham, J., Midden, C.: A robot that says "Bad!": Using negative and positive social feedback from a robotic agent to save energy. In: Conference Proceedings of Human-Robot Interaction, 2009, San Diego, USA (2009)
6. Johnson, D., Gardner, J.A.: The media equation and team formation: Further evidence for experience as a moderator. International Journal of Man-Machine Studies 65, 111–124 (2007)
7. Johnson, D., Gardner, J.A.: Exploring mindlessness as an explanation for the media equation: a study of stereotyping in computer tutorials. Personal and Ubiquitous Computing 13, 151–163 (2009)
8. Johnson, D., Gardner, J.A., Wiles, J.: Experience as a moderator of the media equation: the impact of flattery and praise. Int. J. Hum.-Comput. Stud. 61, 237–258 (2004)
9. Krämer, N.C.: Social Effects of Virtual Assistants. A Review of Empirical Results with Regard to Communication. In: Prendinger, H., Lester, J., Ishizuka, M. (eds.) IVA 2008. LNCS (LNAI), vol. 5208, pp. 507–508. Springer, Heidelberg (2008)
10. Langer, E.J.: Matters of Mind: Mindfulness/Mindlessness in Perspective. Conscious. and Cogn. 1, 289–305 (1992)
11. Midden, C., Ham, J.: Social influence of a persuasive agent: the role of agent embodiment and evaluative feedback. In: Conference Proceedings of Persuasive 2009, Claremont, USA (2009)
12. Nass, C., Moon, Y.: Machines and Mindlessness: Social Responses to Computers. J. of Soc. Issues 56(1), 81–103 (2000)
13. Nass, C., Moon, Y., Carney, P.: Are respondents polite to computers? Social desirability and direct responses to computers. Journal of Applied Social Psychology 29, 1093–1110 (1999)
14. Nowak, K.L., Biocca, F.: The effect of the agency and anthropomorphism on users' sense of telepresence, copresence, and social presence in virtual environments. Presence: Teleoperators and Virtual Environments 12, 2–35 (2003)
15. Reeves, B., Nass, C.: The media equation. Cambridge University Press, New York (1996)
16. Roubroeks, M., Ham, J., Midden, C.: When artificial social agents try to persuade people: The role of social agency on the occurrence of psychological reactance. International Journal of Social Robotics 3, 155–165 (2011)
17. Vossen, S., Ham, J., Midden, C.: What Makes Social Feedback from a Robot Work? Disentangling the Effect of Speech, Physical Appearance and Evaluation. In: Ploug, T., Hasle, P., Oinas-Kukkonen, H. (eds.) PERSUASIVE 2010. LNCS, vol. 6137, pp. 52–57. Springer, Heidelberg (2010)

Tailoring Feedback to Users' Actions in a Persuasive Game for Household Electricity Conservation

Luciano Gamberini[1], Anna Spagnolli[1], Nicola Corradi[1], Giulio Jacucci[2], Giovanni Tusa[3], Topi Mikkola[4], Luca Zamboni[1], and Eve Hoggan[2]

[1] Human Technology Lab, Department of General Psychology, University of Padova
Via Venezia, 8 – 35131, Padova, Italy
[2] HIIT, Department of Computer Science, University of Helsinki, Finland
[3] IES Solutions, Catania, Italy
[4] BaseN, Helsinki, Finland
{luciano.gamberini,anna.spagnolli}@unipd.it,
{nicocorradi,lucazambol}@gmail.com,
{giulio.jacucci,eve.hoggan}@hiit.fi, g.tusa@i4es.it,
topi.mikkola@basen.net

Abstract. Recent work has begun to focus on the use of games as a platform for energy awareness and eco-feedback research. While technical advancements (wireless sensors, fingerprinting) make timely and tailored feedback an objective within easy reach, we argue that taking into account the users' own personal consumption behavior and tailoring feedback accordingly is a key requirement and a harder challenge. We present a first attempt in this direction, EnergyLife, which is designed to support the users' actions and embeds contextualized feedback triggered by specific actions of the user, called 'smart advice'. We conclude by showing the results of a four-month trial with four households that returned promising results on the effectiveness and acceptance of this feature.

Keywords: Sustainability, feedback, adaptive, context aware, design, energy awareness, persuasive technology, smart advice.

1 Introduction

Providing the individual consumer with information about the extent and consequences of his/her own energy consumption activity is considered an effective strategy to foster sustainable consumption practices [4, 5]. Experience accumulated with experiments and field studies shows that generic information is likely to be ignored since people prefer to avoid non-supportive information [2, 12]. Although the most common approach is still to present the user with information and general advice on the total energy consumption, financial measures, or through environmental footprints [10], current technology is also able to collect accurate data and provide the kind of vivid, salient and personal information [3] that is sought by consumers (e.g., [7, 16, 19]). EnergyLife, which is the focus of the present paper, is a mobile application that provides feedback to encourage electricity conservation practices. EnergyLife relies

M. Bang and E.L. Ragnemalm (Eds.): PERSUASIVE 2012, LNCS 7284, pp. 100–111, 2012.

on real consumption data automatically fed into the application by individual electric devices, and returns consumption information along with tips, quizzes, historical data, and a social community. All features are integrated into a simple, usable interface and gradually disclosed according to a game rationale to sustain users' motivation in the long run.

The design of EnergyLife is fully oriented to make its feedback both *action-based and actionable* as a way to apply the principle of tailoring in persuasion (i.e., making persuasive information relevant to the specific characteristics of the recipient [1, 6, 12]) and as a strategy to make feedback more effective. In fact, feedback that helps to identify the relation between users' own actions and a pre-defined goal, e.g. saving electricity, is considered to be more successful [11]. EnergyLife feedback is tailored to the users' actions in three ways, identified through a field trial with an earlier prototype [8]. First of all, by following the users' stepwise increase of awareness when progressively releasing new features of the game. Second, by considering both the individual and the collective (i.e., household) agencies responsible for the game outcome and mapping them differently in the game rationale. Third, by providing tips that are contextualized on the users' consumption behavior (i.e., 'smart advice'). This paper focuses on this last and most original property of EnergyLife, namely the exploration of smart advice tips that are triggered by specific usage patterns and that include customized text. The goal of this paper is to describe the design of these tips, and to report the results of a four-month trial with four households investigating whether this tailoring strategy was well accepted by users and was effective in supporting electricity conservation. We will start by briefly describing EnergyLife and its interface. More specific information about design choices (feedback timing, format, content) can be found in [8, 17]. We will then explain the nature of smart advice tips and the way they were generated. We will then describe the field trial methodology and report the results on users' acceptance and smart advice tips effectiveness. We will summarize our contribution in the conclusion, pointing at possible future developments.

2 A Power Conservation Game in a Nutshell

EnergyLife is an eco-feedback game that provides next-to-real-time, device-based consumption information (i.e., saving and consumption information). The aim of the game is to increase users' awareness of their household's energy consumption.

The main interface consists of a three-dimensional carousel. Each card represents an electrical appliance whose consumption is monitored by sensors (Figure 1a); in addition, a household card represents the overall household and reports data from a sensor installed on the main meter. The fronts of the cards show the current electricity consumption of the device (or the household), and the saving achieved over the last seven days (Figure 1a). The cards can be flipped to access additional information and functionality for the given appliance (Figure 1b), i.e., advice tips, quizzes, and the consumption history for that device (or household) (Figure 2a). The application also offers

information about current levels and scores in the game, and a consumption breakdown per device. Saving information and game scores represent two different kinds of feedback provided by the application, which rely on different agencies, household (i.e., consumption) and individual users (i.e. quiz, messages, …) respectively.

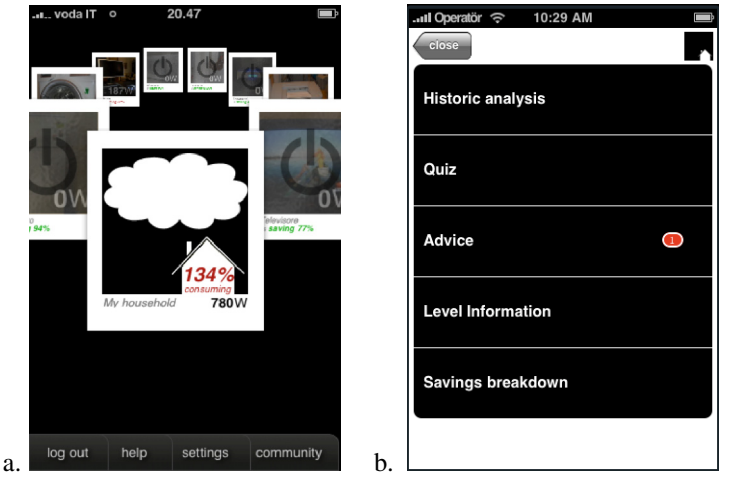
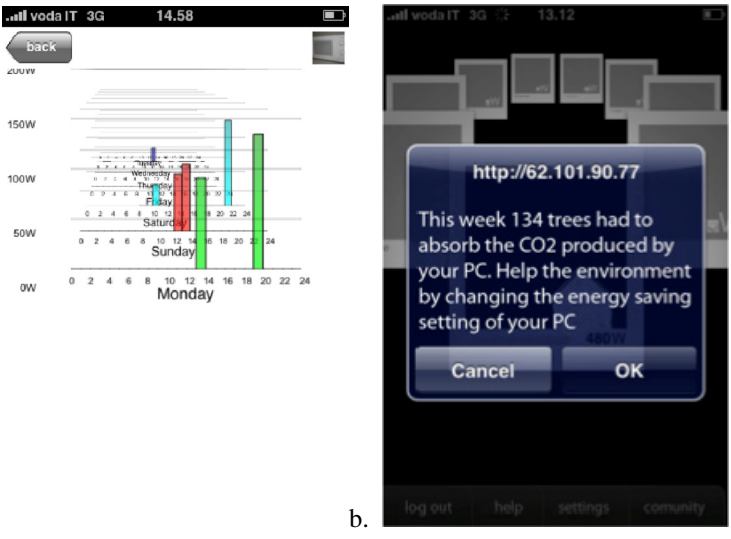


Fig. 1. a. 3D carousel with overview of monitored devices on cards. b. The card is flipped when selected, providing access to a menu.

EnergyLife tries to keep the user enticed and motivated once their initial curiosity drops by implementing game-like features (i.e., scores, quizzes) and by releasing its different features in four steps (i.e., 'levels' in the game) that are designed to fit the subsequent awareness stages of the user [8]. (i) At Level 0 only basic functionality is provided, which consists of saving/consumption feedback compared to a baseline calculated on the previous week. This level allows the users to familiarize with the goal (saving) and with the gap between household consumption and the goal. After 7 days at this level, users are automatically upgraded to Level 1. (ii) In Level 1 the user starts to receive generic advice tips, which direct the user's effort by suggesting energy conservation practices. This represents the knowledge acquisition stage prepared by the previous goal setting stage, and the actual start of the game, since by reading advice tips the individual user acquires points allowing them to advance to the next level. (iii) In Level 2 the user can test his/her knowledge by answering quiz questions that are regularly sent to the mobile phone, and can follow the collection of points in a dedicated page. Users also start to receive smart advice tips in addition to the regular tips activated (see Section 3). By successfully completing the quizzes and continuing to regularly read advice tips, the user progresses to Level 3. (iv) The features in Level 3 allow users to maintain their newly acquired conservation habits via

community sharing and comparative feedback. Users can exchange messages within their own household as well as with users in other households, gaining points for each message, and can see individual and household rankings based on EnergyLife scores.



a. b.

Fig. 2. a. Consumption history of a device b. Visualization of a smart advice tip

Technically, the application client of EnergyLife is a Web application adapted for touch-screen mobile devices. It was developed on HTML5, CSS3 and JavaScript and was deployed on the iPhone 3G and 3Gs. To measure the consumption of appliances in real-time, the system utilizes wireless sensors inserted in plugs. A base-station located in each household caches measurements in non-volatile memory and transfers them to the data storage in bursts. The base station supports both its especially designed 433 MHz wireless network and sensors, but also allows for interfacing with commercial ZigBee based networks. EnergyLife used high quality sensors for the measurements, while the analysis happens in the main computation cloud, for both historical data availability and also software upgradeability reasons – the delay for transferring data to cloud is small with modern IT network capabilities and it gives the system the ability to correlate data form several sources [13]. In a fully realized Internet of Things future, modern devices will contain a microprocessor able to determine the state of the device [18].

3 Smart Advice

Smart advice tips are triggered when exceptional usage of a device is recorded and they provide some tips on ways to avoid such overconsumption in the future. They merge two kinds of information, i.e., consumption information recorded by appliance

sensors, and tips about electricity conservation practices taken from trustable sources (such as environmental associations data). Smart advice tips differ from the regular advice tips provided by EnergyLife in that they are triggered by the specific behavior of the user monitored by the BeAware system, and their content incorporates information on that recorded behavior, along with a tip. In this way, the final smart advice generated is tailored to the specific behavior of the household. Consumption is quantified in kW/h as well as in CO_2, trees compensating CO_2 and even miles on an average city car. For instance, the template "This week 134 trees had to absorb the CO2 produced by your PC. Help the environment by changing the energy saving setting of your PC" involves that a computer was on for longer than a certain amount of time. In the trial, three types of behavior triggered the generation of a related smart advice and its delivery to the phones of the household members:

1. high cumulative usage over a certain period of time: the smart advice was generated after the device was turned on for more than 84 cumulative hours over the last week or had an on-cycle[1] longer than 12 hours or had an on-cycle longer than 5 hours during nighttime;
2. increase compared with past consumption: the smart advice for a given device was generated if its consumption during 2 calendar weeks was higher than the consumption during the previous 2 weeks by 5%, or its consumption during the last 7 days was higher than the consumption during the previous 7 days by 5%;
3. prolonged use of stand-by mode; the smart advice was generated after the device had a stand-by cycle[2] longer than 5 hours or for more than 35 cumulative hours over the last week.

---

[1] As many electrical devices, for example coffee makers and microwaves actually include short power off periods of a few seconds in their normal on-cycle, an on-cycle was defined using the following algorithm:

(I) On-cycle starts when the recorded power has been above a cutoff point of 0.1W for at least 3 consecutive measurements;
(II) (II) On-cycle ends when the recorded power has been below a cutoff point of 0.1W for at least 3 consecutive measurements.

The minimum number of measurements was chosen as a triggering criterion as several common household devices (e.g. microwave oven) exhibit a pattern where a non-maximal level is achieved by shutting down the motor or heating element for a short period of time. Thus the device might have periods of non-activity even when the device as such is in an active state. These measurements mean that the system has to remain on or off for at least 6 seconds to trigger a state change.

[2] Standby mode was detected via the following algorithm:

(I) Detect and identify all stable power states of the device;
(II) If there are more than 2 stable states and at least 1 of them consumes less than 10W, designate the lowest one as stand-by mode.

As the stand-by detection algorithm might fail for devices with particular consumption patterns, its accuracy was confirmed by a human expert for each device. This algorithm was used, as the system capability to fingerprint devices and their states was added relatively late in the project.

The system collected information about the time when the device was used; length of normal usage cycle; energy taken by normal usage cycle; standby pattern; length of standby cycle. This data was compared to all the triggering criteria once a day. If there was a match, the related smart advice was generated (or more than one, if more criteria were met) and sent to household members.

Although we had a larger database of smart advice templates, 13 templates were implemented for the trial, selecting those that related to the devices monitored in all households (i.e., computer, television, refrigerator, washing machine, microwave) and whose triggering conditions did not need any fingerprinting or any additional sensors (e.g. for temperature, light, etc) (Table 1). Each smart advice generated from one template differed from any previous one generated from the same template because it was based on the specific amount of consumption recorded in each case.

Table 1. Templates used to generate smart advice during the trial and number of tips generated

The **computer** that you have left on on [*day*] for *[n]* hours made you consume *[n]* g CO2. Switch off the monitor if you plan not to use it for longer than 15 minutes	50
The **computer** that you left in stand-by for *[n]* hours, made you consume *[n]* kWh this week. Please remember to turn it off completely, to avoid wasting electricity	12
The **computer** that you left in stand-by mode for *[n]* hours this week, made you consume *[n]* g CO2 more than the previous week. Switch the computer off completely instead of leaving it in stand-by	16
This week *[n]* trees had to absorb the CO2 produced by your **PC**. Help the environment by changing the energy saving setting of your PC	30
This week *[n]* trees had to absorb the CO2 produced to provide energy for your **fridge**: you can help the environment by reducing the length at which the door is left open and do not insert food when it is still warm	6
This week your **fridge** spent *[n]* kWh more than last week. To save electricity reduce the duration of door openings and do not insert food when it is still warm	6
This week the **micro wave** oven spent *[n]* kWh more than last week. Please try to use it as little as possible to save electricity	4
This week *[n]* trees had to absorb the CO2 produced by your **micro wave** oven. Save electricity by defrosting your food naturally	12
This week you left the **microwave** on stand-by for *[n]* hours consuming *[n]* kWh. To save electricity, try to turn it off completely	3
The **stereo** that this week you have left in stand-by for *[n]* hours, made you consume *[n]* kWh more than the previous week. Remember to switch it off completely when you don't use it, to conserve electricity	6
On day *[n]* the **TV** left on for *[n]* hours made you consume *[n]* kWh. To save electricity switch it on only when you really watch it	116
On day *[n]* you used the **washing machine** *[n]* times. Was it always fully loaded? In this way you can save electricity and time	86
The day *[n]* you used the **washing machine** *[n]* hours longer than usual. Please try to use it only at full load	150

4 Trial and Evaluation

4.1 Method

A field trial was organized with four households located in urban areas in Catania (Italy), which used the application from January 9th to May 1st 2011. There were ten participants, four women and six men, aged 38 on average (*SD* = 12), and at least two participants belonged to each different household. None of the households included people working in the project. All participants had to own their house and include children since this increased the saving potential; some basic electric appliances (computer, television, refrigerator, washing machine, and possibly microwave) and a Broadband Internet connection was also needed. None had electric heating. Households were recruited through a web form and informal contacts; they were then selected based on an interview and a visit of the house to check the feasibility of the sensing system installation.

Seven sensors were installed in each house (connected to basic appliances listed below plus two devices chosen by the users). All participants were given a smart phone where they inserted their primary SIM card during the trial length, signed a general informed consent form, and an additional consent form at each data collection visit. Data on electricity consumption were collected continuously and automatically during the trial, separately from information on the users' identity. After the first visit when the mobile phones were configured and the users were trained to operate EnergyLife, the experimenters visited the households two more times: at the middle and at the end of the trial, to collect data, and a family interview investigated the users' experience in the trial. We will analyze here the access to the application, a few questions in the satisfaction survey and the household consumption. Some extracts from group interviews conducted during the last visit to the households will be reported, for illustrative purposes only.

4.2 Results

Acceptance and Comprehension

The users kept on accessing EnergyLife over the whole trial duration, gradually reducing the amount of access per day as familiarity with the application and its features increased. Overall access to the application changed (Mann-Kendall τ=-0.432, $p<0.001$) from above 20 login a day at the beginning to about 8 login a day the last day of the trial; the mean number of interactions with devices per login diminished from about 4 per login to 1 per login (Mann-Kendall τ=-0.396, $p<0.001$). This means that once users reach the maintenance phase in which the use of the application as well as the new consumption habits supported by the application got settled, EnergyLife was accessed to reach specific objectives (e.g. checking for new tips, quizzes, and messages).

Among EnergyLife features, the most accessed one was obviously the carousel (4298), followed by consumption history (2167 visits), community (1037), advice (1106) and quizzes (668) (Figure 3). The high number of visits to the consumption history pages shows that users were not only interested in features that were relevant to the game (advice and quizzes) but also in the consumption information that was instrumental to understanding and monitoring their consumption patterns.

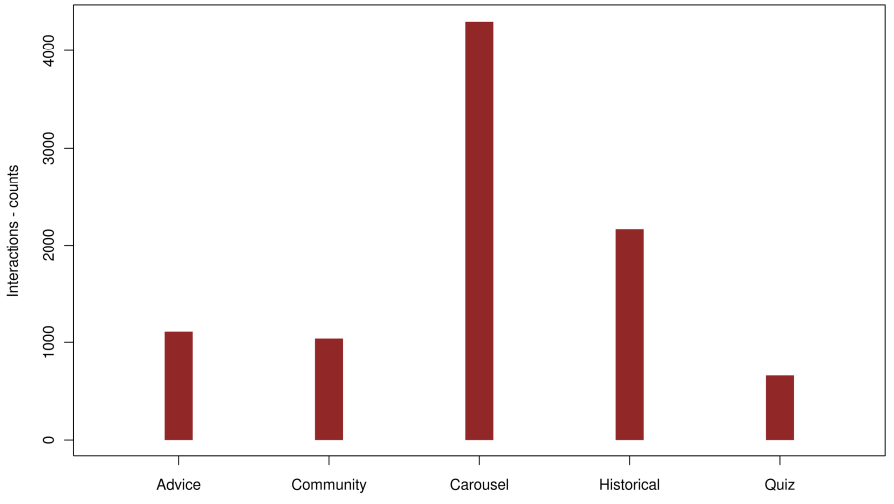


Fig. 3. Number of visits to EnergyLife features during the trial period

A total of 326 smart advice tips were sent and read during the trial (since they occupied the foreground of the screen, the usage of the application was not possible until they were read, see Figure 2b). Each household received an average of 110.24 smart tips (*SD* = 22.35), generated from the same 13 templates. Participants' interest in smart advice was measured on a scale ranging from 1 to 6, where 6 was the highest positive evaluation; the score received was high (mean = 5.00 - first bar in Figure 4). Users also agreed that tips should be (even) more specific, showing their eagerness of receiving detailed information (mean = 4.80 - second bar in Figure 4). Both items were significantly above the middle point of the scale, $t(9) = 2.90$, $p = 0.02$ and $t(9) = 3.34$, $p < 0.01$ respectively. In the interviews users confirmed their interest in the link between their behavior in the house and the comment received in the advice. ".... I wasn't aware about how much CO_2 we produce". We also found that participants were surprisingly neutral about receiving the same kind of tip repeatedly, $t(9) = -0.68$, $p = 0.51$ (mean = 3.20 - third bar in Figure 4). This topic was raised in the interviews, which confirmed that repetitions were noticed but were ultimately considered helpful "…some tips were already applied, but through the suggestions they were strengthened".

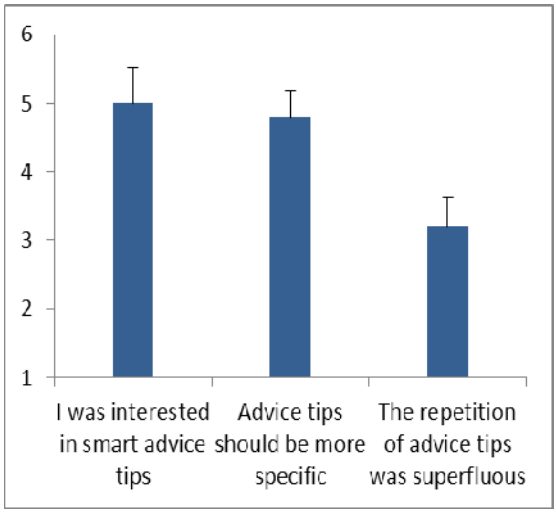


Fig. 4. Answers to four specific items on advice and smart advice tips. Errors bars represent the standard errors.

Participants' comprehension was tested by seven items asking the users to rate the clarity of messages, advice tips and rules from a visual, linguistic and structural point of view. Mean score was 4.80 on a scale ranging from 1 (not at all) to 6 (totally), significantly above the middle point of the scale, thus in the satisfaction range, $t_{(9)} = 5.25$ $p < 0.01$. Learnability was tested with three items such as "*I learned Energy life with some difficulty*", in which users gave positive rates (mean score = 5.13; $t(9) = 7.93$ $p < 0.01$).

Effectiveness of Smart Advice

The average delay between the generation and the reading of a smart advice tip was about 1.5 days, since users did not check the application every day. Therefore, to test the effectiveness of smart advice tips, we compared the electricity consumption of each device the day before and the day after the users read a specific smart advice tip[3]. Since there is not a clear way to determine the identity of the person actually using a device in the analysis, we decided to aggregate the advice per family: the data on smart advice tips reading were aggregated across households. A linear mixed model was used for this comparison. Fixed effects (i.e., the variables actually explored in the analysis) were the consumption time (i.e., the day before or the day after the reading of a smart advice tips) and the trial day. The trial day was included in the model to calculate the effect of the smart advice over and above the global reduction of consumption due to the monitoring per se and other persuasive features of Energy Life. Both household and devices were included as random effects like participants in a repeated measure ANOVA. The results show a significant reduction in the electricity

[3] Therefore, in the case of multiple consecutive smart advice tips about the same device, we compared the consumption of the day before the reading a first smart advice tips with the consumption of the day after the reading of the last tip.

consumption of a device following the presentation of smart advice ($\chi2(1)=6.75$ $p < 0.01$). On average the day after reading of one or more smart advice tips the consumption of the specific devices dropped of the 38% (Figure 5).

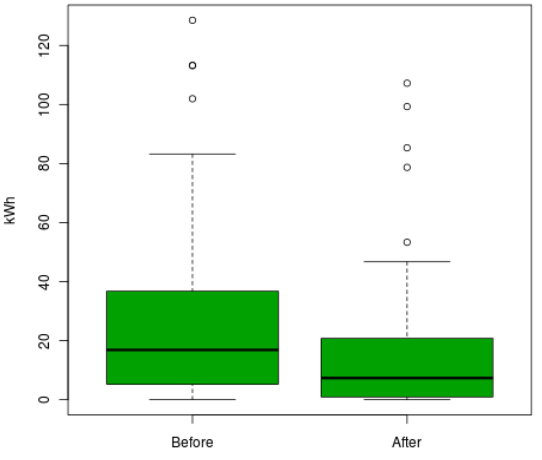


Fig. 5. Box plots of electricity consumptions the day before (left) and after (right) reading a smart advice tip. Isolated dots represent outlier observations.

We also explored the interaction between the effect of a smart advice tip and the trial day (i.e., the variation in the effect of smart advice tips during the trial) with an additional linear mixed model. The interaction was not significant, $\chi2(1) = 0.02$ $p = 0.88$. This result indicates that the effect of the smart advice tips remained stable across the trial, without an effect of habituation or boredom.

Finally, we considered the role of a small series of smart advice tips about the same device read in consecutive days in the same household. This series could unveil a failure of smart advice in reducing consumption. If another smart advice tip was received the next day about the same topic, these series could be due to contingent reasons (i.e., household members did not open the application daily or a high consumption that triggered both a week-related tip and a day-related tip). Participants were considered separately, entered as a random effect in the model together with households, and the consumption the day before and the day after the reading of the first smart advice tip in the series were compared. Again, consumption decreased significantly, $\chi2(1)=10.66$ $p < 0.01$, suggesting that the decrease already reported in Figure 5 is not an artifact of the method adopted to measure consumption at the household level.

5 Discussion and Conclusion

We have described a persuasive eco-feedback game that offers real time, tailored electricity consumption feedback accompanied by tips on energy conservation. This

solution was designed to address crucial challenges, namely making feedback actionable, adapting to the users' varying needs as his/her awareness changed, and maintaining the formation of habits in the long run. We have focused in particular on the role of the smart advice feature, which epitomizes the EnergyLife strategy. Smart advice offer contextualized information and related tips at the device level, whose delivery and content depend on the users' actual consumption behavior. The results of our study indicate that smart advice was well accepted and effective in supporting electricity conservation behavior. Building on these results, the 'smartness' of a tailored, contextualized advice could be improved exponentially to match the accuracy that users expect from a digital sensing and feedback system. They can be based on triggering conditions that take additional information into account (e.g., external temperature, lighting, number of people in the room, automatic recognition of device mode and type of use) to better capture critical users' action patterns.

Acknowledgments. This work was co-funded by the European Union as EU FP7/ICT-2007.6.3, project n. 224557. The authors would like to express their deepest gratitude to the four households participating in the field tests in Catania, to Ilaria Battilani for her help in the data collection and to the other partners in the project (Engineering, Interactive Institute, Enelsì, Vattenfall).

References

1. Abrahamse, W., Steg, L., Vlek, C., Rothengatter, T.: The Effect of Tailored Information, Goal Setting, and Tailored Feedback on Household Energy Use, Energy-Related Behaviors, and Behavioral Antecedents. Journal of Environmental Psychology 27(4), 265–276 (2007)
2. Bem, D.J.: Self-Perception: an Alternative Interpretation of Cognitive Dissonance Phenomena. Psyc. Review 74(3), 183–200 (1967)
3. Costanzo, M., Archer, D., Aronson, E., Pettigrew, T.: Energy Conservation Behavior: The Difficult Path from Information to Action. American Psychologist 41(5), 521–528 (1986)
4. Darby, S.: The effectiveness of feedback on energy consumption. A review for DEFRA of the literature on metering, billing, and direct displays (2006), `http://www.eci.ox.ac.uk/research/energy/downloads/smart-metering-report.pdf` (Last accessed on November 2011)
5. Fischer, C.: Feedback on Household Electricity Consumption: A Tool for Saving Energy? Energy Efficiency 1(1), 79–104 (2008)
6. Fogg, B.J.: Persuasive Technology: Using Computers to Change What We Think and Do. Morgan Kaufmann, San Francisco (2003)
7. Froehlich, J., Findlater, L., Landay, J.: The Design of Eco-Feedback Technology. In: Proceedings of the 28th International Conference on Human Factors in Computing Systems, pp. 1999–2008. ACM, New York (2010)
8. Gamberini, L., Jacucci, G., Spagnolli, A., Corradi, N., Zamboni, L., Perotti, M., Cadenazzi, C., Mandressi, S., Tusa, G., Björkskog, C., Salo, M., Aman, P.: Saving is Fun: Designing a Persuasive Game for Power Conservation. In: Romão, T., Correia, N., Inami, M., Kato, H., Prada, R., Terada, T., Dias, E., Chambel, T. (eds.) Proceedings of the 8th International Conference on Advances in Computer Entertainment Technology, Article 16, 7 pages. ACM, New York (2011)

9. Gustafsson, A., Bång, M., Svahn, M.: Power Explorer: a Casual Game Style for Encouraging Long Term Behavior Change Among Teenagers. In: ACE 2009, vol. 422, pp. 182–189. ACM, New York (2009)
10. He, H.A., Greenberg, S., Huang, E.M.: One Size Does Not Fit All: Applying the Transtheoretical Model to Energy Feedback Technology Design. In: CHI 2010, pp. 927–936. ACM, New York (2010)
11. Kluger, A.N., De Nisi, A.: The Effects of Feedback Interventions on Performance: Historical Review, a Meta-Analysis and a Preliminary Feedback Intervention Theory. Psychological Bulletin 119, 254–284 (1996)
12. Midden, C.J.H., Meter, J.F., Weenig, M.H., Zieverink, H.J.A.: Using Feedback, Reinforcement and Information to Reduce Energy Consumption in Households: a Field-Experiment. Journal of Economic Psychology 3(1), 65–86 (1983)
13. Mikkola, T., Bunn, E., Hurri, P., Jacucci, G., Lehtonen, M., Fitta, M., Biza, S.: Near Real Time Energy Monitoring for End Users: Requirements and Sample Applications. In: Proceedings of IEEE International Conference on Smart Grid Communications (SmartGridComm), pp. 451–456 (2011)
14. Nakajima, T., Vili, L., Tokunaga, E., Kimura, H.: Reflecting Human Behavior To Motivate Desirable Lifestyle. In: DIS 2008, pp. 405–414. ACM, New York (2008)
15. Pierce, J., Fan, C., Llamas, D., Marcu, G., Paulos, E.: Some Considerations on the (In)Effectiveness of Residential Energy Feedback Systems. In: DIS 2010, pp. 244–247. ACM, New York (2010)
16. Strengers, Y.: Designing Eco-Feedback Systems for Everyday Life. In: Proceedings of the 2011 Annual Conference on Human Factors in Computing Systems, pp. 2135–2144. ACM, New York (2011)
17. Spagnolli, A., Corradi, N., Gamberini, L., Hoggan, E., Jacucci, G., Katzeff, C., Broms, L., Jönsson, L.: Ecofeedback on the go: Motivating energy awareness. Computer 44(5), 38–45 (2011)
18. Uckelmann, D., Harrisson, M., Michahelles, F. (eds.): Architecting the Internet of Things. Springer, Berlin (2011)
19. Weiss, M., Mattem, F., Graml, T., Staake, T., Fleisch, E.: Handy Feedback: Connecting Smart Meters with Mobile Phones. In: Proceedings of the 8th International Conference on Mobile and Ubiquitous Multimedia, Article 15, 4 pages. ACM, New York (2009)

Motivational Technologies: A Theoretical Framework for Designing Preventive Health Applications

S. Shyam Sundar[1,2], Saraswathi Bellur[1], and Haiyan Jia[1]

[1] Media Effects Research Laboratory, The Pennsylvania State University, USA
[2] Department of Interaction Science, Sungkyunkwan University, Seoul, Korea
{sss12,sxb979,hjia}@psu.edu

Abstract. Every day, millions of people seek health information online, but we still do not know how to create websites or mobile applications that could motivate them to change their health-related behaviors in a proactive manner. There is a big difference between learning about one's health status and doing something about it. In order to bridge this gap, we provide a theoretical framework for designing *Motivational Technologies*. We discuss how three affordances of modern media interfaces—*navigability*, *interactivity*, and *customization*—could be used to enhance individuals' intrinsic motivation for preventive health, based on self-determination theory. Empirical evidence and design guidelines discussed here could lead to significant advances in health information systems aimed at promoting preventive health behaviors.

Keywords: Motivational technology, preventive health, navigability, interactivity, customization, intrinsic motivation, self-determination.

1 Introduction

Internet-based interactive health technologies, whereby individuals interact with a communication medium to receive and share health information [41], have been credited with transforming healthcare [6, 8]. Websites, online widgets, smartphone and tablet applications make health information more accessible than ever. Recent reports by Mobi Health News [26, 27] identified a total of 5,820 medical, health and fitness apps for smartphones, with an estimated 3 million downloads.

However, despite widespread availability of health applications across several Internet-based media, surveys (e.g., [15]) have shown that most access them only sporadically for information about medication, nutrition and exercise. There is no evidence of sustained use of online health applications that promote and monitor one's preventive health actions. Recent national surveys have shown that very few adults in the United States follow a healthy lifestyle (e.g., [38]), with researchers calling for intensive preventive interventions that will help reduce risk factors for common diseases [22]. Therefore, it appears that, even with available information and tools, most individuals lack the self-determination for preventive health, despite being intrinsically motivated to avoid disease and ill-health. This presents a classic problem for the field of persuasion in general, and persuasive technologies in particular: How

M. Bang and E.L. Ragnemalm (Eds.): PERSUASIVE 2012, LNCS 7284, pp. 112–122, 2012.

can we achieve wider, more consistent use of health technologies in order to reap their benefits for prevention and well-being?

A quick review of recent studies demonstrates several persuasive strategies employed in HCI research. Some have examined how social learning and teamwork principles can be applied to motivate young women to exercise and stay connected via mobile applications [51]. Cell phone applications that keep a count of daily physical activity [9] and caloric balance [52], as well as those that allow users to share health information with peers, friends and families [35], have been tested for personal awareness, competition and social influence. Other studies have explored the role of an embodied conversational agent that can take on the role of an exercise coach and motivate users toward regular physical activity, via goal negotiation and interactive coaching [5]. As we innovate with new tools for serving the cause of preventive health, there is a growing need for assessing the effectiveness of the tools [21]. As a case in point, while download data of mobile apps and evaluations of particular websites are quite promising in their potential to promote diet [33] and exercise [49], we do not yet have systematic knowledge about the efficacy of such tools for preventive health, and what can be done to (i) improve their continued use, and (ii) convert online use into offline health-related outcomes.

Studies have shown that users of Web applications and mobile tools discovered significant design shortcomings, expressing their need for real-time feedback as well as self-configuration and tailoring [2]. A recent study noted that usability issues are likely to frustrate users and lead them to abandon the self-monitoring tasks expected of them [10]. These findings suggest that more interactive, customizable tools and applications may indeed enhance individuals' self-determination to use these online resources for preventive health efforts. In order to meet these needs, we need to articulate and evaluate key guidelines for designing and optimizing such technologies, to introduce system characteristics that promote user engagement and translate into smarter management of a healthy lifestyle, operationalized by sustained performance of wellness routines. We identify key ingredients for internet-based systems that can motivate users to perform health behaviors offline. In this way, our investigation yields guidelines for development of psychologically meaningful interfaces for health websites and mobile applications.

2 Designing Technologies to Enhance Intrinsic Motivation

Intentionality plays a key role in persuasive technologies. Scholars have observed how computers do not have an intention of their own [14, 34]. Instead, intentions lie in the hands of those who design, create and adopt persuasive technologies. With this intentionality perspective in mind, our chief argument is that certain structural features of the technology can be leveraged to build intrinsic motivation among users. For example, scholars have found that the addition of digital photography capabilities to a glucometer changed the device from one that simply displays blood glucose levels to an enabler, encouraging users to visually capture, monitor and regulate their diet [43]. Studies have shown that interactivity in an otherwise static website serves to

profoundly alter user psychology [46]. Investigations into the psychological effects of interactivity and navigability have found that the ability to customize one's online experiences can imbue a profound sense of agency by shifting the locus of control from the system to the user [45].

Understanding oneself is an inherent human drive. Therefore, technologies that showcase one's own self are likely to be intrinsically motivating to users. We can draw upon self-determination theory (SDT) to explain some of the key psychological processes involved in the use of self-monitoring technologies [42]. Self-determination theory makes a distinction between autonomous and controlled regulation of individual behavior. When a goal is set by an individual and is self-driven, then the behavior is said to be autonomous; self-monitoring of one's health-related activities would fall in this category. In contrast, controlled behavior is a reflection of forced adherence to certain goals set by other people (extrinsic factors). Studies on weight-loss, weight-management and diabetes management have provided empirical support to the notion of self-determination and self-regulation by demonstrating that intrinsically motivated individuals are more successful in these programs than those whose behaviors are controlled by external factors. The key challenge therefore is to devise technologies for building self-determination among individuals without the need for repeated and expensive interventions.

2.1 Theoretical Model

Self-determination theory proposes that competence, autonomy and relatedness are essential for one to be intrinsically motivated toward a goal [42]. Autonomy refers to the degree to which individuals feel volitional, i.e., initiators or sources of their own behaviors. Competence is the degree to which one feels able to achieve set goals. The theory also argues that a felt relatedness with one's group, i.e., a warm feeling of connection and sharing, is imperative for enhanced motivation. Individuals who have a stable sense of relatedness with others and who are well integrated in social networks have been shown to possess superior mental and physical health [39]. Likewise, studies have found positive associations between encouraging the need for autonomy and competence in individuals and their health behaviors [28, 31].

Scholars argue that modern day technologies, whether online health-based social networks, bio-medical sensors (e.g., pedometers), personal health dashboards and other health data mining applications have the potential to provide "micro-motivation" by constantly seeking user input and providing continuous feedback that creates an ongoing awareness of their health situation [11]. In line with this, we advocate involving users in their own health behavior activities, letting them navigate, interact and customize their own health website or application. This should produce feelings of competence, relatedness and autonomy respectively, which in turn will positively affect intrinsic motivation to visit the site (or use the app) on a regular basis. Repeated access of health information, especially in such a proactive manner, is a critical determinant of advocated health behaviors, a well-established maxim in the health communication literature [1].

Our theoretical model (Figure 1) predicts not only that navigability (ability to explore the mediated environment), interactivity (ability to interact with others) and customization (ability to tailor the mediated environment) will influence intrinsic motivation to engage with the health tool, but also that this will in turn affect health attitudes (e.g., self-efficacy to learn about and perform preventive health behaviors) and actions (e.g., adoption and maintenance of healthy lifestyle), leading eventually to adoption of preventive health behaviors on a regular basis. Our argument is that specific aspects of health technologies enhance their stickiness, which in turn ensures that users engage these media on a regular basis. Over time, this engagement will translate into sustained exercise of preventive health behaviors. This model may be operationalized in the form of web-based tools or mobile apps for self-management of preventive regimens, such as the ability to monitor one's daily diet in the form of fruit and vegetable consumption, physical activity and exercise routines and allied factors involved in the management of health risk factors. The overarching propositions of the model (Figure 1) are as follows:

P1: Optimal levels of Navigability, Interactivity and Customization will lead to higher levels of Intrinsic Motivation (via psychological mechanisms of Competence, Relatedness and Autonomy, respectively).

P2: Higher levels of Intrinsic Motivation will lead to greater levels of Engagement with health content.

P3: Higher levels of Engagement with health content will lead to better Attitudes and adoption of continuous Health Behaviors.

Most research on health informatics focuses on design and delivery of specific messages for stimulating positive health outcomes. In contrast, our approach explores the role played by technology-related variables in creating an atmosphere conducive to enhancing users' sense of autonomy, competence and relatedness. Thus, our framework provides additional guidelines to build persuasive technologies, which can be easily translated into requirements specification process for software development [34]. In the sections that follow, we introduce the three key design characteristics (navigability, interactivity, and customization) and provide results from empirical studies of these user-interface design variables in a wide variety of content domains.

2.2 Navigability Builds Competence

Navigability refers to the extent to which users can explore a mediated environment, e.g., a health website, in a highly idiosyncratic manner. Online content is organized in a non-linear fashion, with a mix of hyperlinks and other information locatability tools such as search engines. Thus, when seeking information online, users must be able to navigate through large bodies of information. Navigability can be seen as the ability to access any part of an information space [13]. If the structure of a site is very complicated, users will feel lost in a maze of information and feel overwhelmed [17].

A common strategy for ensuring good navigability is to reduce complexity with the help of scaffolds (or help aids) that allow users to explore and manage the

environment. There are two ways in which scaffolding takes place. One is by providing familiar navigational tools such as sitemaps or alphabetical indices that list information in a logical, highly predictable fashion. Another technique is to provide psychologically prominent visual cues (or heuristics) that can help users in processing the information presented to them on health websites. As Balakrishnan and Sundar [3] point out, the primary role of navigability tools is to fulfill a "guidance" or "way-finding" function.

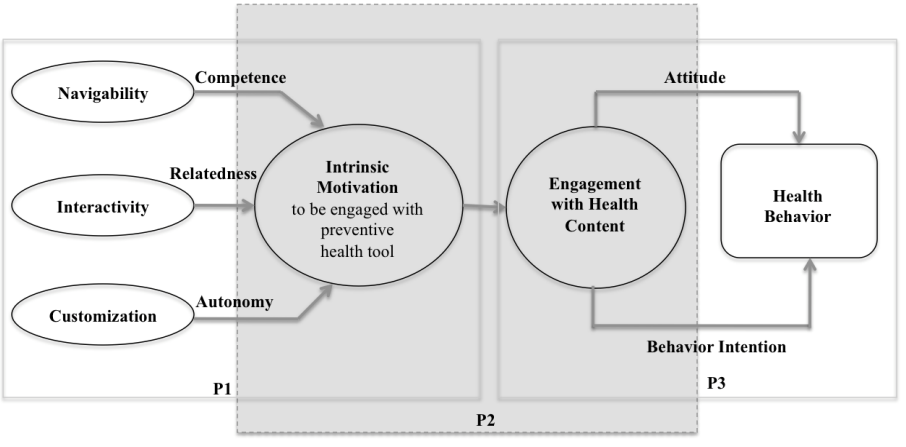


Fig. 1. Theoretical model of motivational technology to promote preventive health behaviors

In sum, effective interfaces should have navigational structures that show users how to reach their goals with ease. This is tied to Information Foraging theory, which suggests that online users are heavily influenced by the "information scent" transmitted by "proximal cues" in the immediate environment about "distal information" [36]. Literature on website usability identifies the following key features for enhancing navigability: search function, drop-down menus, A-Z index lists, breadcrumb navigation, sitemaps and a link to the Home page for user reorientation (see [7, 29, 30]). In addition to usability and user-friendliness, navigability gives the users a sense of control over the website. User control theory [12] suggests that in an online environment filled with clickable links, users will be able to feel in control of content if they are able to navigate through and process content in a manner that they think is personally appropriate. Users must be able to do so at their own pace, and thus any attempt to measure the effects of interface navigability must be based on users' personal needs and abilities. Power users are known to feel a sense of competence when they are able to effortlessly master the quirks of an interface while their lay counterparts end up feeling inadequate when faced with challenging interface features [25, 48]. Sundar [44] discusses the importance of navigability towards building competence and leading to positive cognitive, emotional, and behavioral outcomes by proposing that certain navigation cues on interfaces—cues such as

hyperlinks, site-maps, menus, tabs, and paths in games—serve to trigger distinct heuristics (or mental shortcuts) about the underlying content, leading to quick credibility judgments. Furthermore, navigational design for virtual environments serves the purpose of improving user comprehension of embedded information as well as relationships between different pieces of information [3]. These empirical findings suggest that users' willingness to consume content and be persuaded is significantly affected by the self-efficacy derived from their experience with highly navigable interfaces.

2.3 Interactivity Builds Relatedness

Interactivity, especially "person interactivity" [16], is said to occur if the user is able to connect and have a reciprocal exchange with other people. Good person-interactivity is said to occur when there is a great deal of contingency or threadedness in the interaction [37], i.e., any given message is based on multiple previous messages. This idea of message-based interactivity brings to the fore the notion of relatedness between messages. Studies (e.g., [47]) have shown that higher levels of interactivity will imbue a psychological feeling of relatedness among users and also facilitate greater communal participation in the interaction. When the user of a health site constantly receives messages from other users, it is likely to act as an online support group, providing not only information but also social and emotional comfort, thus building a sense of community. A good example is MoviPill [35], which is a cell-phone application that allows users to monitor not only their own but also their friends' and families' daily medication intake.

Scholars observe that there are two pathways of media influence for promoting changes in health behavior--a direct pathway and a socially mediated pathway--with the latter motivating individuals by connecting them to their social networks and communities [4]. In sum, interactivity features are essential for building relatedness among users as well as a deeper connection between users and the health content that they consume. Sharing options and other forms of interaction between users create a sense of community in online forums or among app user groups. Constructive competition, in workout applications for instance, motivates users in the same way that an offline gym buddy does. On the other hand, regardless of whether interactive features elicit actual interactions, simple interactivity cues, such as bandwagon cues, are likely to engage users by relating to similarly situated others [44], and thereby positively affect health behaviors.

2.4 Customization Builds Autonomy

Customization in health communication has typically taken the form of information tailored in such a way as to meet the needs of one specific individual, based on identification of distinctive attributes of that person [40]. A meta-analysis [32] of 57 studies found that tailoring does have a significant positive effect on health behavior changes. Other research suggests that tailoring may be even more effective when the user does it, especially if s/he is a power-user [48].

According to the agency model [45], customization serves to imbue users with a strong sense of personal agency, by allowing them to constantly specify highly individualistic preferences and requests. In an online environment, the sense of agency is best understood as the degree to which the self feels like s/he is a relevant actor in online interactions, or as a sense of involvement, identity and control, thereby enhancing their efficacy and self-determination in that domain. The fact that the user himself/herself specifies the customizable options makes the locus of causality to be perceived as internal, which brings about the true value of customization, i.e., a real sense of personal autonomy or self-determination [19].

To examine the influence of customization on people's sense of autonomy, we contrasted customization (user-tailoring) with the notion of personalization (where tailoring is performed by the computer system). While system-tailoring results in content that is relevant to the user, customization produces content that is not only relevant but also of utility to the user, thereby boosting user agency and self-determination. Prototypes of several customizable widgets were developed and tested in relation to long-term beneficial health behaviors such as exercise and diet. The findings from this study showed that when users were highly involved in their health, customizing their daily diet and exercise activities enhances their intention to perform preventive health behaviors whereas a control site diminishes their intention over time. Studies on customization features in the Netvibes Web portal [23, 24], avatar customization in Second LifeTM virtual world [20] and self-centered, as opposed to other-centered, customization in iGoogle [18] all show that regardless of the object that they customize, users feel a stronger sense of identity as well as control in their interactions through the interface, compared to users provided with system-tailoring.

3 Implications for Design of Persuasive Health Technologies

To promote sustainable, self-motivating health behaviors, it is not enough to make health information and data available and accessible. Nor is it sufficient to use persuasive messages to convince the user about practicing health behaviors. What we need are technologies that motivate and empower users, not simply inform or persuade them. The key therefore is to move beyond persuasive technologies to what we call *motivational technologies*. While the functional triad of persuasive technologies as tools, media and social actors suggests broad goals for health applications, it is time to drill down and identify specific variables that increase human ability, provide experiences and create relationships in order to persuade as well as motivate [14]. Design elements should be implemented to stimulate the psychology of self-determination for practicing health behaviors in a sustained manner. As discussed above, the technological affordances of navigability, interactivity, and customization have the psychological potential for increasing user's sense of competence, relatedness and autonomy respectively. Therefore, our design suggestions are as follows:

- Interfaces for preventive health applications ought to explicitly factor in competence-building tools. These could be temporary scaffolds and/or visual cues geared to make the user feel competent about his/her ability to navigate the

application and make full use of its features and functionalities. Power users and experienced users ought to be provided shortcuts for improving the efficiency of their actions on the interface.

- Relatedness can be enhanced by a variety of tools, ranging from those that feature message-interactivity (i.e., contingent message exchanges between and among users) to those that offer affordances related to sharing, competing, and otherwise socially exhibiting their preventive health behaviors. In addition, interfaces ought to build in metrics that produce interface cues displaying the degree of social support that users are receiving from their networks for their various health actions.
- For building a sense of autonomy among users, systems ought to offer customization features that actively involve the user in specifying their preferences. This may seem counter-intuitive given the easy availability of personalization technologies that require less user input (because they are able to unobtrusively collate user preferences automatically through their actions online). However, in the interest of building personal agency, it is important for tailoring tools to explicitly offer choices to users and provide them with avenues for creating highly idiosyncratic health regimens. Design should focus on making the user feel like they are the master of their actions, both with the interface and their own bodies. This can be achieved by way of affordances that allow users to exert volitional control (e.g., pursue health goals through an unorthodox patchwork of activities) as well as through interface cues, which highlight the fact that the user is the source of actions triggered by the interface.

For UX researchers evaluating interfaces of health sites and applications, this means

- Devising metrics for assessing perceived competence and related constructs unearthed by our review of studies—constructs such as user control, perceived quality of information scent, wayfinding ease, usefulness of guidance tools, and so on.
- Assessments should include measures of contingency in communications involving the tool (i.e., the degree of threadedness in interactions between the user and specific others), the perceived levels of social support and informational support felt by the user, the degree to which users feel the social presence of their networks, and the extent to which they feel like they can relate with others in making the tool work for them and produce the intended health benefits.
- User studies should factor in the degree of self-tailoring performed by the user, their degree of deviation from defaults and recommended actions, the psychological sense of agency felt by the user, and the extent to which the user sees the interface as reflecting their personal identity.

4 Conclusion

At a recent summit entitled “Putting the I in Health IT,” the U.S. Department of Health and Human Services encouraged health professionals and the general public to think of ways in which health consumers can be empowered and put in charge of their own healthcare [50]. In consonance with such macro-level policy directives, our work builds theoretical knowledge about effects of technological elements in motivating

preventive health behaviors at the individual level. It provides insights into developing and assessing health technologies for their usability and efficacy. Findings from studies employing the proposed model will provide guidelines for designing health technologies to meet the increasing need for visual, emotional and personal applications that utilize mobile and Web 2.0 technologies. Future preventive health technologies should incorporate all three factors discussed here--*navigability*, *interactivity*, and *customization*--in a theoretically meaningful manner so that we can build applications that are usable, social, self-determining, and therefore capable of promoting intrinsic motivation for preventive health actions.

Acknowledgement. This research is supported by the U. S. National Science Foundation (NSF) via Standard Grant No. IIS-0916944 and by the Korea Science and Engineering Foundation under the WCU program funded through the Ministry of Education, Science and Technology, S. Korea (Grant No. R31-2008-000-10062-0).

References

1. Ajzen, I.: The theory of planned behavior. Organizational Behavior and Human Decision Processes 50, 179–211 (1991)
2. Arsand, E., Tufano, J.T., Ralston, J.D., Hjortdahl, P.: Designing mobile dietary management support technologies for people with diabetes. Journal of Telemedicine and Telecare 14(7), 329–332 (2008)
3. Balakrishnan, B., Sundar, S.S.: Where am I? How can I get there? Impact of navigability and narrative transportation on spatial presence. Human Computer Interaction 26, 161–204 (2011)
4. Bandura, B.: Health promotion by social cognitive means. Health Education and Behavior 31, 143–164 (2004)
5. Bickmore, T.W., Caruso, L., Clough-Gorr, K.: Acceptance and usability of a relational agent interface by urban older adults. In: CHI 2005 Extended Abstracts on Human Factors in Computing Systems - CHI 2005, pp. 1212–1215. ACM Press (2005)
6. Brennan, P.F.: Characterizing the use of heath care services delivered via computer networks. Journal of American Medical Informatics Association 2(3), 160–168 (1995)
7. Brower, S.M.: Academic health sciences library website navigation: An analysis of forty-one websites and their navigation tools. Journal of Medical Library Association 92, 412–420 (2004)
8. Combs, S.D.: Startling technologies promise to transform medicine. British Medical Journal 333(7582), 1308–1311 (2006)
9. Consolvo, S., Everitt, K., Smith, I., Landay, J.A.: Design requirements for technologies that encourage physical activity. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems - CHI 2006, pp. 457–466. ACM Press (2006)
10. De Vito Dabbs, A., Myers, B.A., McCurry, K.R., Dunbar-Jacob, J., Hawkins, R.P., Begey, A., Dew, M.A.: User-centered design and interactive health technologies for patients. Comput. Inform. Nurs. 27(3) (2009)
11. Dubberly, H., Mehta, R., Evenson, S., Pangaro, P.: Reframing health to embrace design of our own well-being. Interactions 17, 56–63 (2010)
12. Eveland Jr., W.P., Dunwoody, S.: User control and structural isomorphism or disorientation and cognitive load?: Learning from the Web versus print. Communication Research 28, 48–78 (2001)
13. Fitzpatrick, R.: Additional quality factors for the world wide web. In: The Second World Congress for Software Quality, Yokohama, Japan (2010)

14. Fogg, B.J.: Persuasive technologies. Communications of the ACM 42, 27–29 (1999)
15. Fox, S.: Online Health Search 2006, Pew Internet and American Life Project (2006), http://www.pewinternet.org/PPF/r/190/report_display.asp (retrieved)
16. Hoffman, D.L., Novak, T.P.: Marketing in hypermedia computer-mediated environments: Conceptual foundations. Journal of Marketing 60, 50–68 (1996)
17. Jin, L., Zhu, H., Hall, P.: Adequate testing of hypertext applications. Information and Software Technology 39, 225–234 (1997)
18. Kang, H., Sundar, S.S.: Depleted egos and affirmed selves: The two faces of customization. In: Proc. International Communication Association 2011, Singapore (2011)
19. Katz, I., Assor, A.: When choice motivates and when it does not. Educational Psychology Review 19, 429–442 (2007)
20. Kim, Y., Sundar, S.S.: Me, myself, and my avatar: The effects of avatar on SNW (Social Networking) users' attitudes toward a website and its ad content. In: Wood, N.T., Solomon, M.R. (eds.) Virtual Social Identity and Consumer Behavior, pp. 141–156. M. E. Sharpe (2009)
21. Klasnja, P., Consolvo, S., Pratt, W., Informatics, H., Seattle, I.L.: How to evaluate technologies for health behavior change in HCI research. In: Proc. CHI 2011, pp. 3063–3072. ACM Press (2011)
22. Lloyd-Jones, D.M., et al.: Prediction of lifetime risk for cardiovascular disease by risk factor burden at 50 years of age. Circulation 113, 791–798 (2006)
23. Marathe, S.S.: Investigating the psychology of task-based and presentation-based UI customization. In: Proc. CHI 2009, pp. 3129–3132. ACM Press (2009)
24. Marathe, S., Sundar, S.S.: What drives customization? In: Proc. CHI 2011, pp. 781–790. ACM Press (2011)
25. Marathe, S.S., Sundar, S.S., Bijvank, M.N., et al.: Who are these power users anyway? Building a psychological profile. In: Proc. International Communication Association (2007)
26. Mobi Health News. The fastest growing and most successful health and medical apps (2010), http://mobihealthnews.com/the-fastest-growing-and-most-successful-health-medical-apps/ (retrieved)
27. Mobi Health News. The world of health and medical apps (2010), http://mobihealthnews.com/research/the-world-of-health-and-medical-apps/ (retrieved)
28. Mullan, E., Markland, D.: Variations in self-determination across the stages of change for exercise in adults. Motivation and Emotion 21(4), 349–362 (1997)
29. Nielsen, J.: Site map usability (2008), http://www.useit.com/alertbox/sitemaps.html (retrieved)
30. Nielsen, J.: Breadcrumb navigation increasingly useful (2007), http://www.useit.com/alertbox/breadcrumbs.html (retrieved)
31. Ntoumanis, N.: A self-determination approach to the understanding of motivation in physical education. British Journal of Educational Psychology 71, 225–242 (2001)
32. Noar, S.M., Benac, C.N., Harris, M.S.: Does tailoring matter? Meta-analytic review of tailored print health behavior change interventions. Psychological Bulletin 133, 673–693 (2007)
33. Oenema, A., Brug, J., Lechnher, L.: Web-based tailored nutrition education: Results of a randomized controlled trial. Health Education Research 16(6), 647–660 (2001)
34. Oinas-Kukkonen, H., Harjumaa, M.: A Systematic Framework for Designing and Evaluating Persuasive Systems. In: Oinas-Kukkonen, H., Hasle, P., Harjumaa, M., Segerståhl, K., Øhrstrøm, P. (eds.) PERSUASIVE 2008. LNCS, vol. 5033, pp. 164–176. Springer, Heidelberg (2008)

35. Oliveira, R., Cherubini, M., Oliver, N.: Exploring persuasive techniques for medication compliance. In: Proc. Workshop on Interactive Systems in Healthcare (WISH), pp. 133–136. ACM Press (2010)
36. Pirolli, P.: Exploring and finding information. In: Carroll, J. (ed.) HCI Models, Theories and Frameworks: Toward a Multidisciplinary Science. Morgan Kauffmann Publishers, Francisco (2003)
37. Rafaeli, S.: Interactivity: From new media to communication. In: Hawkins, R., Wiemann, J.M., Pingree, S. (eds.) Advancing Communication Science: Merging Mass and Interpersonal Processes. SAGE Publications, Newbury Park (1988)
38. Reeves, M.J., Rafferty, A.P.: Healthy lifestyle characteristics among adults in the United States, 2000. Archives of Internal Medicine 165, 854–857 (2005)
39. Reis, H.T., Sheldon, K.M., Gable, S.L., Roscoe, J., Ryan, R.M.: Daily well-being: The role of autonomy, competence, and relatedness. Personality and Social Psychology Bulletin 26, 419–435 (2000)
40. Rimer, B.K., Kreuter, M.W.: Advancing tailored health communication: A persuasion and message effects perspective. Journal of Communication 56, S184–S201 (2006)
41. Robinson, T.N., Patrick, K., Eng, T.R., Gustafson, D.: An evidence-based approach to interactive health communication: A challenge to medicine in the information age. Journal of the American Medical Association 280, 1264–1269 (1998)
42. Ryan, R.M., Deci, E.L.: Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. American Psychologist 55, 68–78 (2000)
43. Smith, B.K., Frost, J., Albayrak, M., Sudhakar, R.: Integrating glucometers and digital photography as experience capture tools to enhance patient understanding and communication of diabetes self-management practices. Personal and Ubiquitous Computing 11(4), 273–286 (2007)
44. Sundar, S.S.: The MAIN model: A heuristic approach to understanding technology effects on credibility. In: Metzger, M.J., Flanagin, A.J. (eds.) Digital Media, Youth, and Credibility, The MIT Press, Cambridge (2008)
45. Sundar, S.S.: Self as source: Agency and customization in interactive media. In: Konijn, E., Utz, S., Tanis, M., Barnes, S. (eds.) Mediated Interpersonal Communication. Routledge, New York (2008)
46. Sundar, S.S.: The Oxford handbook of Internet psychology. In: Joinson, A.N., McKenna, K.Y.A., Postmes, T., Reips, U.-D. (eds.) Social Psychology of Interactivity in Human-Website Interaction. Oxford University Press, Oxford (2007)
47. Sundar, S.S., Kalyanaraman, S., Brown, J.: Explicating Web site interactivity: Impression formation effects in political campaign sites. Communication Research 30, 30–59 (2003)
48. Sundar, S.S., Marathe, S.S.: Personalization versus customization: The importance of agency, privacy, and power usage. Human Communication Research 36, 298–322 (2010)
49. Tate, D.F., Wing, R.R., Winett, R.A.: Using Internet technology to deliver a behavioral weight loss program. The Journal of the American Medical Association 285, 1172–1177 (2001)
50. The U.S. Department of Health and Human Services. `http://www.healthit.gov/pledge/?submit.x=135&submit.y=44&submit=Send`
51. Toscos, T., Faber, A., An, S., Gandhi, M.P.: Chick clique. In: Ext. Abstracts CHI 2006, pp. 1873–1878. ACM Press (2006)
52. Tsai, C.C., Lee, G., Raab, F., et al.: Usability and Feasibility of PmEB: A Mobile Phone Application for Monitoring Real Time Caloric Balance. Mobile Networks and Applications 12(2-3), 173–184 (2007)

The Voluntariness of Persuasive Technology

Jilles Smids

Eindhoven University of Technology
P.O. Box 513, 5600 MB, Eindhoven, The Netherlands
j.smids@tue.nl

Abstract. The most important ethical question regarding PTs is the voluntariness of changes they bring about. Coercive technologies control its users by application of direct force or credible threat. Manipulative technologies control their users by influencing them in ways of which the users are not aware and cannot control. As a result, both violate the voluntariness condition of the standard definition of PTs. Any voluntariness assessment needs to consider whether there are external controlling influences and whether the user acts intentionally.

Keywords: voluntary change, persuasive technology, coercion, manipulation, ethics of persuasive technology, controlling influences, intentional action.

1 Introduction

In this paper, it is argued that a voluntary change brought about by a persuasive technology (PT) implies both the absence of controlling influences like manipulation and coercion, and an agent who acts intentional in changing his behavior. As such, the paper is intended as a clarification of the voluntariness clause in the most commonly used definition of PT. This clarification proceeds for an important part through the application of the concept to concrete examples of PTs.

What (Fogg 2003, 1,15,16) writes about the definition of PT can be combined as follows:

PTs are technologies which are intentionally designed to change the behavior, attitude or both (without using coercion or deception; persuasion implies voluntary change).

This definition has become more or less the standard definition, which is evident from for example (IJsselsteijn et al. 2006, 1) and (Oinas-Kukkonen 2010, 6). Oinas-Kukkonen defines a behavior change support system, a type of PT, as "an information system designed to form, alter, or reinforce attitudes, behaviors or an act of complying without using deception, coercion or inducements[1]". He goes on to state that "persuasion relies on the user's voluntary participation in the persuasion process".

[1] By inducements he means economic incentives. Because these seem less important in the field of PT, they are not considered in this paper.

M. Bang and E.L. Ragnemalm (Eds.): PERSUASIVE 2012, LNCS 7284, pp. 123–132, 2012.

All three sources mentioned recognize the evident importance of separating persuasion from coercion and deception (a form of manipulation), and fortunately at the same time all three of them show awareness of the fact that PTs cannot be made voluntary by defining them to be so: "the line between persuasion and coercion can be a fine one" (Fogg 2003, 21) and "...more subtle and effective persuasive techniques....will continue to pose new moral dilemmas." (IJsselsteijn et al. 2006, 4) and "Quite surprisingly, ethical considerations have remained largely unaddressed in persuasive technology research. Many important issues need to be recognized, such as the actual voluntariness for change..." (Oinas-Kukkonen 2010, 10).

Ethical reflection on PT should consider intentions of the persuaders, behavioral and attitudinal aims of the PT, and methods of persuasion (Berdichevsky and Neuenschwander 1999; Fogg 2003). However, reflection on methods and the voluntariness for change they imply is by far most important and most urgent for two reasons. First, as autonomy and freedom are fundamental values in western societies, users of PT have a prima facie basic right not to be influenced in ways that violate voluntariness, even not in cases where the intentions behind the PT and its aims are praiseworthy. *Ends* simply do not immediately justify their *means*; additional justification is always necessary. For example, it can be argued that road safety justifies technical enforcement of speed limits in order to prevent harm to others. Health however seems not to justify technology that manipulates people to behave healthier, because mainly harm to self is at stake. Second, if PT turns out not to rely on voluntary change, it ceases to be PT *by definition* and should no longer be called, sold, and promoted as such.

As little reflection on the voluntariness of PT has been done so far, in this paper the actual voluntariness of examples of (types of) PTs will be discussed, and some conceptual clarification of 'voluntariness' will be given. In sections 2 and 3 examples of technologies which are coercive and manipulative are discussed. Section 4 contains a discussion of two potential conditions for voluntariness: intentional action on the part of the agent and substantial freedom from controlling influences. In the final section some conclusions are drawn and suggestions for future research are given. Throughout the paper some recommendations for PT designers are given and some research questions for PT scholars are specified.

2 Coercive Technologies

This section will give a short characterization of coercion which subsequently is used to discuss the extent to which some examples and types of PT are in fact coercive technologies. In this manner, it should become clear in which ways coercive technologies violate the voluntariness condition of the standard definition. Coercion can be brought about either by means of direct force or by (sometimes implicitly) making a credible threat. Direct force is most often thought of as physical force by which the agent is guided into performing the action intended by the coercer. An example would be a piece of coercive technology that, applying satellite-based localization, limits the speed of a car to the legal maximum at that location. As will be argued below, the type of direct force which can be applied by some PTs will be more likely of the type that proceeds by using up human self-regulatory capacity to such extent that the user

finally will give in and change his behavior (attitudes in such cases usually remain unchanged and the behavior often will be counter-attitudinal).

In case of coercion by threat, the coercer communicates to the coercee that he will bring about some consequence, which is credible and so undesirable to the coercee that she performs the action the coercer wants her to. There has been some discussion over the question whether the threatened person actually must give in for there to be coercion, or whether coercive intentions and actions on the part of the coercer are sufficient for coercion (Anderson 2011). This is relevant to PT, because not every user will be affected in the same way by a given PT that applies (slightly) coercive methods; one may be coerced, another may resist.

An example of what is sometimes presented as PT (Spahn 2011) but which is in fact coercive technology are the well-known blinking lights and obtrusive sounds that operate in case a driver does not fasten the seat-belts. Clearly, for some drivers it is not by voluntary choice that they fasten their seatbelts, but they only do so because they cannot bear the extremely annoying sounds. Had the car not been equipped with the device, they would not have chosen to fasten their seat-belts. The type of coercion at hand can be related to both credible threats and to direct force. If a driver initially does not fasten his seat-belts, but finally gives in because he no longer is able to put up with the highly irritating noise, he is directly forced by the noise. If he initially fastens his seatbelt in order to prevent the noise, he does so because of the credible (i.e. certain) threat of the noise. Either way, what is decisive here is that the noise is irritating to such a great extent, that for most people it is psychologically impossible to bear, or at least far more than sufficiently unattractive. In order to change this piece of coercive technology into a genuine PT, the lights and sounds should stop operating after a few seconds, such that it only reminds the driver, but subsequently leaves him free in how to act upon the reminder.

This example is a good illustration of one of the advantages of PTs over human persuaders mentioned by (Fogg 2003): PTs are persistent. In this case however, persistence leads to coercion, especially because the device that makes the sound 'controls the interactive possibilities' (Fogg 2003, 216), i.e., the driver cannot turn of the device. A further advantage of PTs that can result in coercion is their potential for being ubiquitous. If a PT combines all of those characteristics, the power-relation between PT and user will be asymmetric, sometimes to the extent that the technology becomes coercive.

This is especially the case due to our limited human capabilities and resources for self-control. Research done in the last decades, e.g. by (Muraven and Baumeister 2000), shows that exercising self-control consumes resources which are limited. Exercising self-control for some span of time turns out to result in a significantly lessened degree of self-control in a subsequent task which requires self-regulation. So, persistent attempts at technological persuasion may be resisted by the user up to the point where even the strong-willed ones have exhausted their resources available for self-control. In such cases, the resulting change will be clearly involuntary, because it is contrary to the previous attempts by the user not to change and not to give in. In other words, the genuine will of the user is expressed by her resistance, where her giving in is just the result of exhausted self-control. This process of deliberately causing the using up of the resources needed for self-control can appropriately be identified as the use of direct force and thus as coercion.

If another characteristic of PT mentioned by Fogg is added, namely its ability to manage and make use of huge volumes of data, then the (at least theoretical) possibility of abuse of our limited capacities for self-control emerges. PTs can acquire personal persuasion histories, resulting in personalized 'persuasion profiles', which contain information about the types of influences an individual is especially susceptible to and under what circumstances (Kaptein and Eckles 2010). The most powerful and (probably coercive) PT that could be imagined therefore is a PT which knows our persuasion profile, which can attempt to persuade us at any time and at any place, and which is persistent and cannot be turned off.

Designers of PTs should try to avoid such types of technological coercion.[2] Because the potential persistence of PTs threatens their very nature as PT (as this implies voluntary change), designers should provide means for the users to turn the PT (temporarily) off. This possibility at the same time enables users to control the places where they wish to be available for attempts at technological persuasion. The phenomenon of persuasive profiling needs much more ethical reflection than can be pursued here. Probably legal restrictions on acquiring, using, and selling information on personal profiles will be needed and appropriate.

In the next section, a similar discussion will be given for manipulative technologies as was given in this section for coercive technologies.

3 Manipulative Technologies

Perhaps PTs run a greater danger to become manipulative than coercive, because the boundary between persuasion and manipulation might be less clear. A neutral and technical meaning of the word manipulation is 'handling', 'manoeuvring' (Brenkert 2008, 155–6). In the case of objects, this can be achieved by physical means, but in the case of humans this would amount to coercion. In order to bypass human resistance in handling them (thus treating the person as an object), the manipulator has to hide the fact that he is manipulating the person in some way. Therefore, the principle underlying manipulating is controlling persons in such way that their awareness of the very fact that they are manipulated is prevented. When this is done by giving people false or incomplete information, we speak of deception.

The first question that can be asked is whether so-called ambient PTs allow for voluntary change. Ambient PTs aim to influence through mechanisms that require little cognitive resources and little focal attention of the recipient. (Ham and Midden 2010) gave participants in a study the goals to regulate a thermostat such that optimal comfort is reached and that as little energy as possible is used. It turned out that giving ambient feedback through lighting the wall resulted in the lowest energy use compared to giving factual feedback through numbers. Although more research is needed, it seems that this lighting feedback has its persuasive potential not through conscious cognitive elaboration, but through pathways of which a user is less aware.[3]

---

[2] See (P.J. Nickel and Spahn, A., submitted) for an interesting discussion of how to make the human-PT relation less asymmetric.

[3] These pathways could be described using the language of the well-known elaboration likelihood model of persuasion.

Does this lack of awareness make PT using lighting feedback pieces of manipulative technology? Not necessarily so. Most importantly, the participants were given an explicit goal, using less energy, and they knew the corresponding aim of the lighting feedback and also the meanings of the colors green and red (low versus high energy consumption). If, translated to a real life setting, users have the opportunity to turn off the feedback, then, because they know the persuasive aim, the users can make an informed an voluntary decision on whether to use it or not. That opportunity in itself does not make the methods of PT voluntariness preserving. So the important and to my knowledge not yet answered research question is: how does the effect of the lighting feedback depend on whether the user shares the goal of energy reduction or not? For, if he does not share this goal but very much prefers comfort, and the lighting feedback nonetheless is still effective, than he is influenced in a direction against his will. If on the other hand he does share the goal, then his change in behavior is in line with his general intention to reduce his energy consumption. Whether the change was voluntary seems therefore still an open question. In any case, research by psychologist PT researchers on the influence of goal-sharing on the effectiveness of ambient PT is welcome.

Another type of PT that is potentially manipulative is technology that enables so called 'unconscious persuasion' (Ruijten, Midden, and Ham 2011.). The authors show that research participants who are primed with the goal 'to perform well', performed better under influence of subliminal feedback compared to the condition of no feedback. Importantly, in this study the goal was primed, i.e., was created in such way that the participants were not aware of their having this goal. Of course, this need not be the case; subliminal feedback on consciously self-chosen goals is also possible. Still, 'unconscious persuasion' is an oxymoron, for by definition it does violate the voluntariness condition, for it exactly fits the description of manipulation given above: a person is under control of the manipulator outside his awareness.

One might object that the case of ambient PT is not different: persons are not aware of the way the lighting feedback is influencing them and they do not exert conscious control over its workings. In addition, if the use of subliminal feedback is disclosed, including its aim, are there any differences left with ambient PT? The most important difference is that in ambient persuasion the user is always in a position to recognize *that* he is being influenced. And if he turns to a reflective mode, he can, for example, think about the meanings of the green and red color and reassess whether energy use reduction is an aim he wants to pursue. So, each instance of attempts of ambient persuasion can be recognized by a user. In case of 'unconscious persuasion' however, even if its application is disclosed to him, he cannot recognize an instance of it, because the subliminally presented messages are not consciously recognizable.[4] Consequently, he cannot recognize *that* he receives subliminal feedback and also cannot know *what* it's content is.

It is not without reason that subliminal advertizing is prohibited and PT scholars should not label this technique persuasion, for this will result in stretching this term far too much and also in inflation of and loss of public trust in the concept of PT. Ruijten et al. show awareness to these issues, but still hold that 'unconscious persuasion' "… can be considered as a useful approach that has high potential persuasive

[4] See for this distinction (Bovens 2009).

power, provided that sufficient care is given to an ethical design". This might all be true, but ethical design cannot change subliminal feedback into persuasion. Again, there may be situations in which the use of subliminal feedback can be ethically justified, but it simply should not be called persuasion, just in order to protect the 'brand of PT'[5].

In the next section a conceptual clarification of 'voluntariness' will be provided and linked to the above discussions of coercive and manipulative technologies.

4 The Concept of Voluntariness

The idea of a voluntary change of a person's behavior or attitude has a number of different but related connotations. A voluntary change can be thought of as a change which is freely made, intentional, purposeful, unforced, or the result of a conscious act of the will. The Latin origin of the word is 'voluntārius' which stems from 'voluntas', and ultimately from 'velle', meaning to will, to wish. ("Free Merriam-Webster Dictionary (Voluntary)".). For purposes of assessing the voluntariness of change under influence of a PT, a more systematic, conceptual analysis of 'voluntariness' is needed.

One place to look for such analysis is the current debate in the bioethics literature on the condition of voluntariness as one of the elements of informed consent. (Nelson et al. 2011) argue that an action is voluntary if and only if the action is i) intentional (the agent is *in control*) and ii) substantially free of controlling influences (*non-control* of the agent).[6] For example, if one walks to the fridge because one wants a beer, one acts intentionally. If one is prevented from taking the beer by her roommate who blocks her way to the fridge, then one is subject to a controlling influence, in this case coercion; as a result one does not act voluntarily. Nelson's analysis will be taken as starting point and both conditions will be discussed, starting with the second. Subsequently, this discussion will be related to the role played by the voluntariness condition in the standard definition and to technological manipulation and coercion.

Nelson et al. discuss three external controlling forms of influence: coercion (either by force or by a severe and credible threat), manipulation (e.g. withholding information, misleading exaggeration) and undue inducement (e.g. strong economic incentives).[7] Interestingly, they see persuasion as the "paradigm of an influence that is both non-controlling and warranted". And "[i]t is in part warranted *because* it is non-controlling". It must be said, however, that they define persuasion explicitly as rational persuasion, in which the "merits of reasons proposed" do the work of belief change. This might well be taken as an overly narrow definition of persuasion and in any case,

---

[5] (P. J. Nickel, manuscript.) makes this point.

[6] Both conditions have a longstanding history, including Aristotle in Bk 3 of his Nichomachean Ethics.

[7] The authors also discuss internal controlling conditions which prevent actions from being voluntary, such as influences caused by certain forms of mental illness (e.g. psychopathy, addiction). These conditions are not discussed in this paper because the influence of PT is by definition external. Still, those internal controlling conditions will be relevant in cases of many PTs designed for health and well-being, because such PTs often intend to support people who lack sufficient self-control, e.g. in matters of healthy eating.

it does not square with the practice of PT research and applications. For, this practice shows a focus that is wider than belief change and in which many other sources of influence (e.g. social praise, emotion, and social cooperation) are studied in addition to the influence of reasons proposed by the persuader. Consequently, persuasion will not be a non-controlling form of influence by definition and careful study of actual instances of PT and their influence strategies is necessary.

Importantly, condition (ii) speaks of 'substantial freedom' from controlling influences. The authors argue that control and non-control come in degrees. However, they "[…] do not attempt to determine precisely the degree of control an agent must have over causal influences in order to act voluntarily…"(p10). This is regrettable, for this is a crucial question, although admittedly a difficult one. Still, without knowing the exact degree of control over external influences an agent must have, it seems natural to hold that an increase in such control signifies an increase in voluntariness and vice versa. Applied to PT, the more manipulative or coercive its methods are, the less voluntariness is left to its users. As a consequence, it becomes less appropriate to call such technology PT and additional justification is needed before it could be employed.

Regarding condition (i), of intentional action, Nelson and co-authors hold that "[w]e perform intentional actions when we act in accordance with plans proposed for the execution of an action". Furthermore, they hold that "[a]cts are either intentional or nonintentional. While there may be gray cases, such as acting on 'automatic pilot,' having a confused plan of action, or having a minimal rather than an elaborate plan of action, even these cases can be classified as either intentional or nonintentional."

Acting according to a plan may sound as acting consciously according to a deliberate decision. Although matters are complicated here, on such specification of intentional action, much everyday behavior would not be voluntary. One of the major developments in social psychology during the last decades concerns the growing insight in the major role of automaticity in human behavior[8]. From these findings about the role of automaticity in the origin of our behavior, it follows that often we do not act according to a clear, conscious, and deliberate plan. But in many cases we would nonetheless like to say that we act voluntarily, so either intentionality is not a necessary condition for voluntariness, or we need to analyze the concept of intentionality in terms which make less reference to consciousness, reflectivity, and purposefulness.[9]

No matter what exactly the authors using the standard definition mean themselves, it is important to get clear on how to interpret the voluntariness phrase with respect to this intentional action condition. For, on the above strong interpretation of intentional action, a large classe of current PT might not count as relying on voluntary change. Take for example the ambient PTs which, as discussed above, aim to influence through mechanisms that require little cognitive resources and little focal attention of the recipient. Prima facie, it is most natural to see the change in attitude and / or

---

[8] See e.g. (Bargh and Chartrand 1999).

[9] A third possibility is of course that we act voluntarily much less often than we thought and that our judgment of those many ordinary cases in which we act unreflectively in an more or less automatic fashion as voluntary actions is incorrect; see (Blumenthal-Barby 2011, 25). From such perspective, it might be more appropriate to say that such actions are neither voluntary, nor involuntary.

behavior resulting from the use of ambient PT as not according to an explicit an consciously pursued plan. At the same time however, as argued, ambient PTs do not immediately seem to rule out voluntariness.

The observation that many actions which are not according to a clear and consciously pursued plan are nonetheless clearly voluntary, suggest that a more moderate interpretation of intentional action is preferable. This can be illustrated by analyzing the phenomenon of habits. Habits are formed by automation processes due to repeated identical actions under the same circumstances. Now, if one acts completely out of habit and without any plan of which one is aware and without any clear and conscious intention, one still can act voluntary. For, one's habitual behavior is the expression of deeply ingrained goals one has which in certain circumstances are automatically activated and which lead automatically to behavior. Take for example going to your work. Just the very general and implicit intention to go to your work leads to the performance of a whole range of actions, e.g. taking your bike to travel the station, checking in for the train, checking out, etc. Each action contributes to the overall intention of going to your work and as such is an intentional action, but which can proceed without much or even without any deliberation and conscious attention.

Applied to PT this means that types of PT that engage our less reflective and less cognitive processes do not immediately have as a result that the user is not in control anymore. To return to the class of ambient PTs once more, *if* these are successful in changing the behavior of users because of their having a (general) intention to perform the target behavior, then the change in behavior counts as voluntary. So, if a user has a general intention to lower her energy use and if that is what causes ambient PT using lighting feedback to be effective, than her behavior change is an intentional action and thus voluntary. More research is needed however into the precise psychological mechanisms of those types of persuasion in order to solve that issue.

Regarding PT that is potentially coercive, the relevant issue is not whether there is a clear intention to act or not. Coercion is a form of external controlling influence. Both when a coercer proceeds by making credible threats and by applying direct force, the user is aware of what is going on. In some sense, the resulting change is an intentional action; one intentionally fastens one's seat-belt in order to prevent the obtrusive sound. However, this intention originated *due to coercion* and therefore isn't the expression of an agent *in control*.

From the discussion of this section, it can be concluded that in assessing the voluntariness of PT, it is needed to investigate both into potential circumstances of external controlling influences (manipulation and coercion) and into the condition intentional action. Both conditions, provided interpreted properly, are individually necessary and jointly sufficient conditions for voluntary action.

5 Conclusion and Suggestions for Further Study

In this paper it was first briefly argued that the most important ethical question regarding PTs is the voluntariness of changes they bring about. Subsequently the ways in which coercive and manipulative technologies control their users and as a result violate the voluntariness condition of the standard definition of PT were discussed. Any voluntariness assessment needs to consider whether there are external controlling

influences and whether the user acts intentionally. This condition of intentional action should not be interpreted as necessarily involving consciousness, deliberativeness, and purposefulness.

It is recommended that designers of PT always perform a voluntariness assessment of the PT they create. This is essential if they in fact want to design persuasive technology and if they want to prevent that the public comes to distrust the idea of PT. The fact that all the three authors discussed in the introduction show a clear awareness of the fact that some PTs might rely on less voluntary change seems to be a clear indication that designers and PT scholars should be capable of doing so.

Some issues for further study can be identified. First, voluntariness was mainly treated as an all-or nothing concept, whereas voluntariness comes of course in degrees. Second, the voluntariness for change will be different for different users, due to their differences in capacities, attitudes, interests, etc. Any adequate treatment of these issues need more conceptual analysis (ethicists) and more study of the precise psychological mechanisms and processes at work in users of PTs. This latter study should include the question to what extent the persuasive potential of ambient forms of PT, which do less engage our reflective system, depends on the measure in which users share the persuasive goals of the PT. As a result of this further research, ethicists and PT scholars should be better equipped to determine the degree of voluntariness enabled by PT.[10]

References

1. Anderson, S.: Coercion. In: Zalta, E.N. (ed.) The Stanford Encyclopedia of Philosophy (Winter 2011), http://plato.stanford.edu/archives/win2011/entries/coercion/
2. Aristotle, Nichomachean Ethics, trans. Ross, W.D., http://classics.mit.edu/Aristotle/nicomachaen.html (accessed from March 9, 2012)
3. Bargh, J.A., Chartrand, T.L.: The unbearable automaticity of being. American Psychologist 54(7), 462–479 (1999)
4. Berdichevsky, D., Neuenschwander, E.: Toward an ethics of persuasive technology. Commun. ACM 42(5), 51–58 (1999)
5. Blumenthal-Barby, J.S.: On the Concept and Measure of Voluntariness: Insights from Behavioral Economics and Cognitive Science. The American Journal of Bioethics 11(8), 25–26 (2011)
6. Bovens, L.: The Ethics of Nudge. In: Grüne-Yanoff, T., Ove Hansson, S. (eds.) Preference Change, pp. 207–219. Springer Netherlands, Dordrecht (2009)
7. Brenkert, G.G.: Marketing Ethics. Wiley-Blackwell (2008)
8. Fogg, B.: Persuasive technology: using computers to change what we think and do. Morgan Kaufmann Publishers, Amsterdam (2003)
9. Free Merriam-Webster Dictionary (Voluntary), http://www.merriam-webster.com/dictionary/voluntary (accessed on December 23, 2011)

[10] Philip Nickel, Andreas Spahn, and Frank Verberne are acknowledged for many helpful discussions. Three anonymous reviewers are thanked for their helpful comments and suggestions.

10. Ham, J., Midden, C.: Ambient Persuasive Technology Needs Little Cognitive Effort: The Differential Effects of Cognitive Load on Lighting Feedback versus Factual Feedback. In: Ploug, T., Hasle, P., Oinas-Kukkonen, H. (eds.) PERSUASIVE 2010. LNCS, vol. 6137, pp. 132–142. Springer, Heidelberg (2010)
11. IJsselsteijn, W.A., de Kort, Y.A.W., Midden, C., Eggen, B., van den Hoven, E.: Persuasive Technology for Human Well-Being: Setting the Scene. In: IJsselsteijn, W.A., de Kort, Y.A.W., Midden, C., Eggen, B., van den Hoven, E. (eds.) PERSUASIVE 2006. LNCS, vol. 3962, pp. 1–5. Springer, Heidelberg (2006)
12. Kaptein, M., Eckles, D.: Selecting Effective Means to Any End: Futures and Ethics of Persuasion Profiling. In: Ploug, T., Hasle, P., Oinas-Kukkonen, H. (eds.) PERSUASIVE 2010. LNCS, vol. 6137, pp. 82–93. Springer, Heidelberg (2010)
13. Muraven, M., Baumeister, R.F.: Self-regulation and depletion of limited resources: Does self-control resemble a muscle? Psychological Bulletin 126(2), 247–259 (2000)
14. Nelson, R.M., et al.: The Concept of Voluntary Consent. The American Journal of Bioethics 11(8), 6–16 (2011)
15. Nickel, P.J.: The Definition of Persuasive Technology (manuscript)
16. Nickel, P.J., Spahn, A.: Design for Symmetry - Trust, Discourse Ethics, and Persuasive Technology (submitted)
17. Oinas-Kukkonen, H.: Behavior Change Support Systems: A Research Model and Agenda. In: Ploug, T., Hasle, P., Oinas-Kukkonen, H. (eds.) PERSUASIVE 2010. LNCS, vol. 6137, pp. 4–14. Springer, Heidelberg (2010)
18. Ruijten, P., Midden, C., Ham, J.: Unconscious Persuasion Needs Goal Striving. The Effect of Goal Activation on the Persuasive Power of Subliminal Feedback. Paper presented at Persuasive 2011, Columbus, OH, USA (2011)
19. Spahn, A.: And Lead Us (Not) into Persuasion...? Persuasive Technology and the Ethics of Communication. Science and Engineering Ethics (Published online first May 5, 2011), doi: 10.1007/s11948-011-9278-y

Persuasion and Reflective Learning: Closing the Feedback Loop

Lars Müller, Verónica Rivera-Pelayo, and Stephan Heuer

FZI Research Center of Information Technology, Karlsruhe, Germany

Abstract. Reflecting about past experiences can lead to new insights and changes in behavior that are similar to the goals of persuasive technology. This paper compares both research directions by examining the underlying feedback loops. Persuasive technology aims at reinforcing clearly defined behaviors to achieve measurable goals and therefore focuses on the optimal form of feedback to the user. Reflective learning aims at establishing goals and insights. Hence, the design of tools is mainly concerned with providing the right data to trigger a reflection process. In summary, both approaches differ mainly in the amount of guidance and this opens up a design space between reflective learning and persuasive computing. Both approaches may learn from each other and can use common capturing technologies. However, tools for reflective learning require additional concepts and cues to account for the unpredictability of relevance of captured data.

1 Introduction

Encouraging reflection or directly persuading a person can induce a change in behavior. Reflective learning and persuasive technology both investigate potential support for this changes by using information technology. The research field of persuasive computing aims at influencing human behavior by computational means. Gathering data about behavior to enforce desired behavior and point to possible flaws is a common method, known as reinforcement. Tools like UbiFit Garden [3] monitor behavior like the physical activity and provide feedback to motivate more physical activity. Learning by reflection can support the definition of goals, changes in perspectives and ultimately changes in behavior, as well. Here, an overlap between both concepts has been identified that has not been sufficiently explored.

Both kind of applications use feedback loops to support behavioral change. Feedback loops have been a topic of research since Rosenblueth et al. [29] published "Behavior, Purpose and Teleology" and laid the foundation for cybernetics. If a system uses its own output to control its behavior a feedback loop is created. Feedback loops have been used to model and analyze complex systems in many research fields, i.a. cybernetics, biology, climate and social sciences.

Capturing of behavior, e.g. by sensors, plays a major role to establish these feedback loops in persuasive technology and reflective learning. Sensors support the self awareness of a user by providing data about current behavior, including

M. Bang and E.L. Ragnemalm (Eds.): PERSUASIVE 2012, LNCS 7284, pp. 133–144, 2012.

trends and patterns [13]. Current persuasive technology solutions focus on seemingly straight forward changes in behavior that turn out to be very difficult to achieve and need additional support, e.g. doing more sports [3,23,12] or minmizing one's carbon footprint [11]. The user is supported by encouraging simple and clear defined task to reach long term goals [1]. This clear definition of desired behavior guides the selection of useful sensors and capturing tools, but there are more complex goals, that cannot be associated with single and clear defined tasks, e.g. optimizing work performance or changing social behavior.

According to Boud et al. [2], learning by reflection (or *reflective learning*) supports learning by returning to and evaluating past experiences in order to improve future experiences and promote continuous learning. However, the existing theories on reflective learning were developed at a time, where the hand written diary was the only supportive tool. Today, sensors and information systems capture and provide much more data, that is already used in many persuasive applications. Although the research in computer supported reflective learning lacks behind persuasive technology, the underlying theories have been refined for more than eighty years.

This paper analyzes the approaches from reflective learning and persuasive computing and sheds light on the underlying feedback loops. Can both fields of research learn from each other, e.g. by adapting concepts and reusing existing technology? This paper reviews the underlying assumptions and theory as well as the current technology. The main focus is capturing technology, which is successfully used by persuasive computing and allows capturing large amounts of data needed for reflective learning in a wide range of application domains.

The next section introduces the theory and application examples behind reflective learning and persuasive technology that target behavioral change. Drawing from these fundamentals, the two different approaches and their consequences for tool development are compared by mapping them into a general feedback loop. Afterwards, we review existing persuasive products and technologies according to their applicability for reflective learning and the ethical implications. The conclusion summarizes the contribution and provides a brief outlook on further research.

2 Background

Research in reflective learning and persuasive technology has created a plethora of theories on reflection and a wide range of persuasive tools that tackle different aspects of changing behavior. This section describes the two approaches and their application.

2.1 Reflective Learning

Decades of research in reflective learning have highlighted different aspects of reflective learning, leading to multiple theories [6,15,2,30]. Hence, it is difficult to define a shared understanding of reflection. In the following, the most important

approaches are briefly summarized. A more detailed description and discussion of existing approaches can be found in [20].

Nearly all research on reflective learning refers back to the idea of experiential learning by Dewey [6], i.e. we learn by comparing our expectations to what we experience. Our expectations form a continuity that is built on former experiences. We learn by adapting this continuity in interaction with the environment or as Dewey puts it: "Continuity and interaction in their active union with each other provide the measure to educative significance and value of the experience". This can be understood as a description of a feedback loop. The learner adapts to the environment by learning. However, Dewey's work is mainly concerned with the benefits of reflective thinking for formal learning, hence, emotional aspects are rather neglected.

Kolb [15] describes experiential learning in the form of a cyclic process; the so called Kolb Cycle. Reflective observation is one of four components of the Kolb cycle. The cultivation of the capacity to reflect in action (while doing something) and on action (after you have done it) has become an important feature of professional training programs in many disciplines [30]. One example building on the approach of reflective practice is presented by Daudelin, who defines reflection as "the process of stepping back from an experience to ponder, carefully and persistently, its meaning to the self through the development of inferences; learning is the creation of meaning from past or current events that serves as a guide for future behavior" [5].

There is a vast body of research, but this research does not take sufficiently into account the possibilities provided by information technology. Even newer work like Moon [20] and Daudelin [5] use only traditional instruments like learning journals and structured interviews. Therefore, we a theory that provides insights into the cognitive processes and can be a basis for the integration of technology into the reflection process. We chose the model introduced by Boud et al. [2] as theory for this paper because it considers the complete cognitive process, including affective aspects, but does not define specific activities around this process or a specific domain. In the model, reflective learning refers to "those intellectual and affective activities in which individuals engage to explore their experiences in order to lead to new understandings and appreciations" [2]. Therefore, the reflective process is based on the experiences of the learner, which are considered as "the total response of a person to a situation, including behavior, ideas and feelings". The process described by Boud et al. consists of three stages, in which the learner re-evaluates past experiences by attending to its various aspects, and thereby producing outcomes. The defined outcomes can be cognitive, affective or behavioral. The reflection process and it's context, experiences and outcomes, are depicted in Figure 1.

The start of the reflection process is a critical point for tool support that initiates the return to experiences. Boud et al. do not explicitly define the beginning of the reflection process because "most events which precipitate reflection arise out of normal occurrences of one's life." However, the provided examples can be easily linked to cognitive dissonance theory [7]. Cognitive dissonance theory

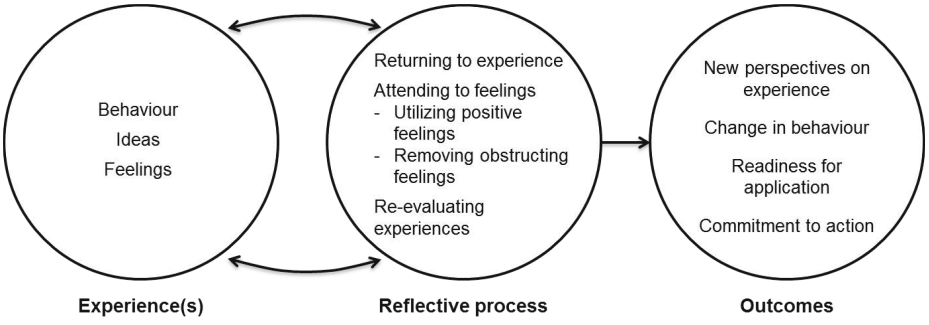


Fig. 1. The reflection process in context [2]

describes how a mismatch between attitudes and behavior could lead to rethinking of attitudes and experiences.

Current research in the MIRROR project [19] or by Li [16] points towards a growing integration of information technology into reflective learning. Driven by the fast development of applications like RescueTime [28] that support a continous capturing of data for reflection, research is now responding. RescueTime captures the amount of time an application is used at the PC and thus supports reflection about time and task management. These reflection applications rarely reference the above mentioned theories but use the described cycles or feedback loops to trigger reflective thinking and induce behavioral change.

2.2 Persuasive Technology

While reflective tool support is still its infancy, persuasive technology spans a wide field of applications and and approaches. The scope of this paper are applications and research that is inspired by conditioning technology which Fogg [9] defines as "a computerized system that uses principles of operant conditioning to change behavior". Operant conditioning refers to the work of Skinner [32] who created a feedback loop by providing strong reinforcement.

Many persuasive tools [3,1,14] are using feedback loops to trigger and support behavioral changes. These tools implement persuasion strategies as described by Fogg, i.e. simplifying of tracking and reinforcement. Both strategies are employing an underlying feedback loop, by mirroring back recorded behavior. The simplification of tracking leaves the interpretation of the data to the user, but by capturing the right data, e.g. plotting sports activities vs. weight, a certain behavior can be sustained.

Mobile applications play a major role in many persuasive approaches, because they can establish a closer connection to a person; they do not wait until a user asks for feedback but can provide it at any time [12]. Therefore, the feedback loop does not suffer from a delay. Fogg [9] refers to this perfect point in time to act as Kairos factor. In the moment a behavior is performed, a message about this behavior is more likely to have an impact.

Current persuasive computing solutions focus on health and fitness applications, where strict adherence to rules results in clearly measurably progress, such as weight loss. Ubifit [3] is an example of a persuasive tool that uses reinforcement. It aims to facilitate physical actitivity by displaying the activity – measured by an acceleration sensor and biosensors – on a Smartphone. The main focus of research is the design of the appropriate presentation and reaction to this data, using a display of a virtual garden that rewards the user if he does sufficient activity (showing more flowers on it).

Today, more and more products are using these concepts to induce and support behavioral change. *Up* is a recent product of Jawbone [14] that has received a lot of attention. It uses a wrist worn acceleration sensor and a website to visualize the data and collect additional feedback. It simplifies the tracking of activity and can vibrate to remember the user of activity. The website offers additional feedback, e.g. users can join a challenge.

More overt reinforcement is used by beeminder.com. Users have to select a goal and can pay an amount of money to help staying on track. In practice, every time the planned progress is not reached some of the money is deduced. If the final goal is reached, the user will get all the money back. This is an obvious kind of negative reinforcement. While this is a very strong motivation, this kind of feedback might actually cause people to quit using the tool.

Hence, research in persuasive technology places a strong emphasis on the design of the feedback. For example, Consolvo et al. [4] provide a guideline for the development of persuasive applications, whose first point lists "abstract & reflective" as a key attribute of a successful implementation beneath attributes like aesthetic, unobstrusive, public and controllable. This hints towards an overlap between persuasive technology and learning by reflection.

Despite the growing list of products and research prototypes, current apps and even research projects are mainly successfull in a limited number of application domains, e.g. healthcare and sustainability, where progress and success can be easily measured. Selecting the right behavior that can be measured and sustained is a critical step in the design process [8]. However, for many application domains there are not yet sufficient capturing concepts available, because the goals are rather fuzzy and/or only cues can be captured that leave to much room for interpretation. In this case, the feedback loop cannot be closed.

3 Persuade or Support Reflection

This section analyzes reflective learning tools and persuasive tools in respect to their design goals, their methods and implementations by mapping them into a common model based on a feedback loop.

Reflective Learning and persuasion by reinforcement are both aiming at changing behavior by implementing a feedback loop that is increasingly based on technology. Figure 2 shows the feedback loop established between cognitive processes, the resulting behavior and the tools that measure behavior and mirror it back to the user. We have used the model by Boud et al. to describe the cognitive

processes of a user. As explained above, this model provides a broadly applicable model that does not limit the used tools or the contents of reflection. Two feedback loops are shown:

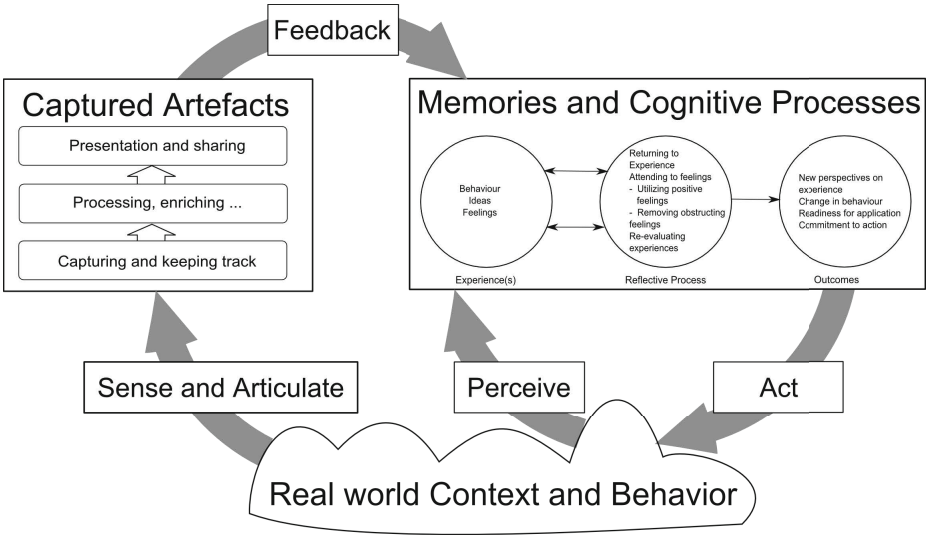


Fig. 2. The underlying feedback cycle of reflective learning and persuasion

(a) The first loop comprises the user, the real world context and her behavior. This is the normal flow of action and reactions of a user without tool support. The user acts in the environment, perceives the effect and thus creates new experiences. The result are new memories and, in some cases, a discrepancy will trigger a reflection process, which in turn may result in a change in behavior. This is the approach as described by Dewey [6].

(b)The second loop includes captured artefacts that provide an alternative version of the real world context and behavior and are captured through articulating of perceived experiences and sensing of context data. This digital version of the experience is limited by the capturing methods and the chosen presentation. The captured artefacts can support reflective learning or directly persuade towards a predefined goal. Therefore, reflective learning as well as persuasive applications capture data, process and enrich it before they present it to the user or share it with other users. The resulting feedback provides an alternative version of a behavior and it's context that the user might have perceived or remembers completely different.

3.1 A Matter of Guidance?

The design of the feedback decides if an application supports reflection or directly persuades a user to induce behavioral change. According to Boud, reflective learning requires an active process of returning to experience, attending to

feelings and re-evaluating the experience. The reflective process should result in an outcome that could be a change in behavior. Therefore, reflective learning tools aim at showing discrepancies between the own mental model and the recorded data. On the contrary, persuasive tools work in a limited domain and aim at giving specific advice to change a behavior. This advice is infered from the recorded data and specific to the targeted behavior. In summary, persuasive tools provide more guidance to change the behavior. They reduce the cognitive effort for the user to change the behavior. Reflective learning aims at increasing the cognitive effort by raising awareness [18,16,28] for an issue and trigger additional cognitive processes.

When analyzing the existing persuasion tools, this distinction begins to blur. While persuasive approaches use mainly positive and sometimes negative reinforcement, this categorization is difficult for many of them. For instance, Jawbone Up [14], is mainly reminding the user to create an awareness. Only a few applications like beeminder.com [1] use strong reinforcement, while more and more applications are simplifying the capturing and quantifying behavior to raise awareness about one's own behavior. The Quantified Self Community [33] is one of the drivers of this current trend.

3.2 Different Requirements for Capturing

These differing goals of the feedback for persuasion and reflection lead to different challenges in the capturing of behavior. While persuasive computing needs clear indicators to provide a very specific advice, reflective learning has to deal with larger amounts of data. One of the main challenges for reflective learning tools is to show the discrepancy between the memories of an event and what really happened, because the only proof of what really happened is captured by sensors and tools which may be –intentionnally (by design) but also unitentionally – biased as well or capture only a subset of the event. In result, each event has three representations: (a) what really happened, (b) what the user believes has happened and (c) what tools have captured about that event.

The real event (a) is composed of a large number of components, ranging from things like time and place to body temperature and minor details of the surrounding that have an absolut precision. Tools and humans can capture only a subset of these details with a limited precision.

A person experiencing the same event (b) will perceive only a subset of all details. For example in a fierce argument, the facial expression of another person may remain in the memory but not all the details about the conversation before the incident or even the physiological signals of the other person leading to this reaction. Moreover, this limited information is interpreted by cognitive processes and adapted to the mental model. The result is often a limited and sometimes skewed version of the real event, because memory is malleable [17]. It is a reconstructive process that can be influenced by a person's mental model [24] and by post-event information [34].

Tools and sensors record a third version of the event (c). Sensors are often specialized on single details like temperature, location or the heart rate of a user.

However, current systems aim at reconstructing higher level properties that are more useful to a user. Other sensors, like cameras, capture more information than we can analyze. Image recognition therefore focusses on compressing the image to filter relevant details. Nevertheless, the limitations of the recording and the achieved precision differs from what really happened. This difference may act as a cue for users to remember events, discover discrepancies, start to reflect and maybe change their behavior. Moreover, Sellen and Whittaker [31] argue that we should rather focus on capturing memory cues instead of a full fledged life logging of all activities and details.

Persuasive tools often rely on a single simple sensor that is very effective to measure a certain behavior. This helps to minimize the effort for the user and the danger of misinterpretation of the data. If a tool would show obviously wrong data to the user, it would loose its persuasive power. If this approach is extended to more complex application domains, multiple sensors will become necessary that can be only partially linked to a specific behavior. Each new sensor can provide access to complete new ways to understand and change our behavior. Reflective learning tools minimize the neccessary interpretation by leaving the interpretation to the user. The clear specific goal is replaced by the general goal of more awareness about own behavior and the flaws of our mental models and cognitive processes.

4 Closing the Feedback Loop

Capturing data about behavior is a crucial step in reflective learning and persuasive technology applications to provide the appropriate feedback at the right point in time. Data that describes behavior can be split into two categories: experential data and quantitative data. Experential data relates to the subjective experience of an event or situation. The captured digital representation of an event is therefore very similiar to the memory of the user. In contrast, quantitative data intends to be as objective as possible and the resulting representation of an event is much closer to reality. While experential data may yield deeper insights about behavior and it's reasons, quantitative data is easier to analyze by algorithms and can be gathered automatically by sensors. Hence, quantitative data is currently driving persuasive technology and reflective learning.

While new persuasive strategies and user interface approaches have substantially transformed applications to be more persuasive, the most advanced capturing mechanisms [3,22] are now several years old. Current persuasive applications focus on simple capturing concepts that have a strong relation to goals and relevant behavior. Fogg [8] explicitly recommends to limit the complexity and start with a small part of the target behavior that is easy to measure and probable to change. Reflective learning on the other hand requires a broader amount of information to allow open reflection. Therefore, several capturing solutions have to be combined to capture relevant behavior. Existing capturing concepts in persuasive computing provide a good starting point to select sensors and tools that capture relevant behavior, but the broad range of possible outcomes of the reflective process raises new requirements.

4.1 Capturing Quantitative Data about Behavior

The reuse of sensors and technology to gather data about the behavior of a user can reduce the required effort to create new reflective learning tools. Moreover, even existing and already deployed persuasive technology products might be integrated. The used capturing concepts can be roughly categorized by their technology.

The majority of persuasive tools relies on active user reporting to capture accurate high level data and can be used in a wide range of domains. However, the benefit for capturing this data has to be clear, otherwise the motivation of the user will decline before an actual change in behavior is achieved. The strategies of existing products to persuade users to capture data at regular intervals range from the hedonistic value of the application to short-term rewards for capturing data. The MoodMap App [21] leverages the curiosity of the user. The user is rewarded for each input of a current mood with the average mood of the whole group. The most widely used sensor is the acceleration sensor that can capture body movement if worn at the wrist or in a pocket. This movement is used to analyze general physical activity, count the number of steps or monitor movement at night to estimate the sleep quality. Acceleration sensors are widely available in smartphones, but it remains unclear where it is worn on the body. Specialized sensors, as used by Nike+ [23], have known location and thus their data can be analyzed with a better confidence in the results.

Biosensors are increasingly used to measure actvity and health at a more specific level. Morris [22] used a mobile ECG (Electrocardiography) to monitor heart deseases and the MAHI project [18] used a glucose meter and a phone to support diabetes patients with. These projects aim at very specific illnesses, but biosensors can be used as well to estimate the affective state of a user [27]. This new application domain is currently growing and pointing towards new application domains for persuasive technology, e.g. emotional awareness. Another upcoming topic for persuasive technology [11,26] are tools that support the reduction of water and energy consumption. Water or energy meters are used as sensors to measure the current consumption and offer hints to increase awareness and induce a change in behavior.

4.2 Unpredictability of Relevance of the Captured Data

Reflective learning tools are not geared towards enforcing a specific behavior but aim at creating awareness and a cognitive dissonance to identify possible improvements. Therefore, reflective learning tools have to capture a wider range of data to account for the multiple possible outcomes. However, only a fraction of this data can trigger a reflective process and should be presented to a user. A selection is required to reduce the amount of relevant data.

The captured sensor data is probably too complex to infer interesting patterns for reflection, but some sensors offer cues to identify relevant time spans and thus filter the data. Three different kinds of capturing cues are promising candidates to judge the relevance: (a) arousal and other emotional cues, (b) social interaction and (c) tasks and activities.

(a) Critical incidents are prone to trigger emotional reactions such as arousal that can be measured to a certain extend by psychophysiological sensors, e.g. Affectiva Q [27].

(b) Reflection topics are often related to other persons and interaction, e.g. disagreements between collegues, working together or serving a customer. Although, computer mediated communication is already recorded by many tools, face to face interaction plays a major role. These contacts can be captured by reality mining technologies like the SociometricBadge [25].

(c) A link to currently performed tasks and activities would allow the identfication of patterns across multiple performances of the same task. If patterns are identified, deviations from normal operations can be detected. Tasks can be captured by augmenting tools, e.g. company information systems, or diary like applications, e.g. MyExperience [10].

All of these three cues can be used to aggregate data and create a layer of abstraction that can provide new insights. At this abstract level, sharing between different users becomes possible because tasks can be compared as well as reactions to similar situations.

4.3 Ethical Considerations

Persuasive technology and reflective learning are similar in their use of feedback loops. Both capture behavior of a user to create and adapt feedback to influence this user. How this feedback is used is the main design difference between both approaches. The strong goal focus of persuasive applications results in a strong guidance while reflective learning leaves much more responsibility to the user.

This may lead to the notion that reflective learning would deal with less ethical issues than persuasive technology because the user has more control and has to conciously decide to change her behavior instead of being nudged into a direction that might be against her own good. However, reflective learning requires more captured data which on the one hand may interfere with privacy standards and on the other hand provides a more complete feedback that competes with the perceived version of an experience. This may alter our view on experiences on a much deeper level as we expect. This data may even influence our identity as it shapes our autobiographic memory. Unknown biases in the capturing or the presentation may unintentionally influence users. Hence, huge ethical implications have to be taken into account in the design of reflective learning applications so to avoid rumination but facilitate effective reflection.

5 Conclusion

This paper has analyzed the different approaches of persuasive technology and reflective learning applications. While both approaches come from different theoretical backgrounds, a closer look reveals a common design space between them that could be used to create new tools to achieve behavioral change in a wider number of domains. Reflection is a promising approach to achieve behavioral change in domains where relevant behavior is difficult to measure. Both approaches build on different theories but they share similiar goals and methods

that can be understood by looking at the underlying feedback loops. The differences lie in the amount of guidance and the richness of captured data.

Reflective learning applications can integrate or build on existing capturing solutions from persuasive technology, but they have to deal with the unpredictability of relevance of data. Three types of cues have been proposed to filter and connect the different data sources in order to estimate the relevance of captured experiences and identify critical intervention points: emotional cues, social interaction and task/activity tracking.

More research is necessary to explore the specifics of the overlap between persuasive and reflective applications and to integrate current trends like The Quantified Self [33]. There may as well be combinations between both aproaches, such as applications that persuade users to reflect on complex topics that cannot be tackled by persuasive technology.

Acknowledgements. This work has been co-funded by the European Commission within the 7th Framework Programme in the MIRROR project (http://www.mirror-project.eu).

References

1. beeminder.com (2012), http://www.beeminder.com
2. Boud, D., Keogh, R., Walker, D.: Promoting Reflection in Learning: a Model. In: Reflection: Turning Experience into Learning, pp. 18–40. Routledge Falmer, New York (1985)
3. Consolvo, S., Landay, J.A., McDonald, D.W.: Designing for behavior change in everyday life. Computer 42, 86–89 (2009)
4. Consolvo, S., McDonald, D., Landay, J.: Theory-driven design strategies for technologies that support behavior change in everyday life. In: CHI 2009, pp. 405–414. ACM (2009)
5. Daudelin, M.W.: Learning from experience through reflection. Organizational Dynamics 24(3), 36–48 (1996)
6. Dewey, J.: Experience and Education. Macmillan, London & New York (1938)
7. Festinger, L.: A theory of cognitive dissonance. Stanford Univ. Press (1957)
8. Fogg, B.J.: Creating Persuasive Technologies: An Eight-Step Design Process. In: Persuasive 2009. ACM (2009)
9. Fogg, B.: Persuasive technology Using Computers to Change What We Think and Do. Morgan Kaufmann Publishers (2003)
10. Froehlich, J., Chen, M.Y., Consolvo, S., Harrison, B., Landay, J.A.: MyExperience: a system for in situ tracing and capturing of user feedback on mobile phones. In: MobiSys 2007, pp. 57–70. ACM, New York (2007)
11. Froehlich, J., Dillahunt, T., Klasnja, P., Mankoff, J., Consolvo, S., Harrison, B., Landay, J.A.: Ubigreen: investigating a mobile tool for tracking and supporting green transportation habits. In: CHI 2009, pp. 1043–1052. ACM, New York (2009)
12. Gasser, R., Brodbeck, D., Degen, M., Luthiger, J., Wyss, R., Reichlin, S.: Persuasiveness of a Mobile Lifestyle Coaching Application Using Social Facilitation. In: IJsselsteijn, W.A., de Kort, Y.A.W., Midden, C., Eggen, B., van den Hoven, E. (eds.) PERSUASIVE 2006. LNCS, vol. 3962, pp. 27–38. Springer, Heidelberg (2006)

13. Goetz, T.: Harnessing the power of feedback loops (June 2011), http://www.wired.com/magazine/2011/06/ff_feedbackloop/all/1
14. Jawbone Up (2012), http://www.jawbone.com/up
15. Kolb, D.A.: Experiential Learning: Experience as the source of learning and development. Prentice Hall, Englewood Cliffs (1984)
16. Li, I., Dey, A.K., Forlizzi, J.: Understanding my data, myself: supporting self-reflection with ubicomp technologies. In: UbiComp 2011, pp. 405–414 (2011)
17. Loftus, E.F.: Planting misinformation in the human mind: A 30-year investigation of the malleability of memory. Learning & Memory 12, 361–366 (2005)
18. Mamykina, L., Mynatt, E., Davidson, P.: MAHI: investigation of social scaffolding for reflective thinking in diabetes management. In: Proceedings CHI 2008, pp. 477–486 (2008)
19. Mirror - reflective learning at work (2012), http://www.mirror-project.eu
20. Moon, J.A.: Reflection in learning and professional development. Routledge (1999)
21. Mora, S., Rivera-Pelayo, V., Müller, L.: Supporting Mood Awareness in Collaborative Settings. In: CollaborateCom 2011 (2011)
22. Morris, M.E.: Technologies for Heart and Mind: New directions in embedded Assessment. Intel. Technology Journal 11(01) (2007)
23. Nike+ (2012), http://www.nikeplus.com
24. Norman, D.A.: Explorations in Cognition. W.H.Freeman & Co Ltd. (1974)
25. Olguin, D., Waber, B.N., Kim, T., Mohan, A., Ara, K., Pentland, A.S.: Sensible organizations: Technology and methodology for automatically measuring organizational behavior. IEEE Transactions on Systems, Man, and Cybernetics-Part B: Cybernetics 39(01) (2009)
26. Patel, S.N., Robertson, T., Kientz, J.A., Reynolds, M.S., Abowd, G.D.: At the Flick of a Switch: Detecting and Classifying Unique Electrical Events on the Residential Power Line (Nominated for the Best Paper Award). In: Krumm, J., Abowd, G.D., Seneviratne, A., Strang, T. (eds.) UbiComp 2007. LNCS, vol. 4717, pp. 271–288. Springer, Heidelberg (2007)
27. Poh, M., Swenson, N., Picard, R.: A wearable sensor for unobtrusive, long-term assessment of electrodermal activity. IEEE Transactions on Biomedical Engineering 57(5), 1243–1252 (2010)
28. Rescuetime (2012), http://www.rescuetime.com
29. Rosenblueth, A., Wiener, N., Bigelow, J.: Behavior, Purpose and Teleology. Philosophy of Science 10(1), 18–24 (1943)
30. Schön, D.A.: Educating the Reflective Practitioner. Jossey-Bass (1987)
31. Sellen, A.J., Whittaker, S.: Beyond total capture: a constructive critique of lifelogging. Commun. ACM 53, 70–77 (2010)
32. Skinner, B.F.: About behaviorism. Knopf (1974)
33. The Quantified Self (2012), http://quantifiedself.com
34. Wright, D.B., Loftus, E.F.: Eyewitness memory, pp. 18–40. Routledge, New York (2008)

Normative Social Influence in Persuasive Technology: Intensity versus Effectiveness

Thijs Waardenburg[1,2], Robbert Winkel[1,3], and Maarten H. Lamers[1]

[1] Media Technology group, LIACS, Leiden University, The Netherlands
[2] Utrecht University of Applied Sciences, Utrecht, The Netherlands
[3] The Hague University of Applied Sciences, The Hague, The Netherlands
thijs.waardenburg@hu.nl, r.n.w.winkel@hhs.nl, lamers@liacs.nl

Abstract. It has been established that normative social influence can be used effectively in persuasive technology. However, it is unknown whether the application of more social pressure makes it more effective. To test this hypothesis, a quantitative experiment was conducted on the online social network Facebook. Although evidence to support the hypothesis was found, it cannot be concluded from this experiment that more intense persuasion is more effective, when utilizing normative social influence in persuasive technology.

Keywords: Persuasive technology, captology, normative social influence, mass interpersonal persuasion, social networks.

1 Introduction

Every person is influenced by his or her social environment, in particular by groups he or she belongs to such as family, colleagues or teams [1]. This is generally known as social influence; "[the] process whereby attitudes and behaviour are influenced by real or implied presence of other people" [2, p.236].

Deutsch and Gerard distinguish two processes of social influence: 'informational social influence' and 'normative social influence' [3]. The former can be defined as social influence based on the need to know what is right when we cannot form our own opinion, and seek assurance by looking for evidence or perceived expertise of other persons. This research, however, is focused on normative social influence, which is based on the need to be accepted within a group, potentially in conflict with our own values, intentions or opinions. Psychologists call this 'peer pressure' or 'pressure to conform' to the norms of a group [1, p.199].

Persuasion is a form of social influence that Zimbardo and Leippe define as "changing a person's behaviours, feelings or thoughts about an issue, object or action" [4]. Fogg defines it as "an attempt to change attitudes or behaviours or both (without using coercion or deception)" [5, p.15]. Fogg clarifies that persuasion is restricted to voluntary change: a villain threatening you with a gun and telling you to give him all your money is *coercive* rather than persuasive, (a gun would otherwise be an example of rather potent persuasive technology). The definition for persuasion that

M. Bang and E.L. Ragnemalm (Eds.): PERSUASIVE 2012, LNCS 7284, pp. 145–156, 2012.

will be used for this research is therefore: "The attempt to change behaviours, feelings or thoughts about an issue, object or action, without using coercion or deception".

In 1997 a group of researchers at the CHI97 conference established an area of research that explores persuasion by means of technology [6]. This research area, called persuasive technology or captology (Computers As Persuasive Technology), focuses on what Fogg defines as "any interactive computer system designed to change people's attitudes or behaviours" [5, p.1]. Since then, research into the possibilities of persuasive technology has been carried out into a wide range of fields, including public safety, health-care and education [7–9]. A new form of persuasive technology, called Mass Interpersonal Persuasion (MIP) which focuses on persuading large groups of people in an online social environment, emerged in 2008 [10]. It uses group dynamics, and social acceptance and rejection to be effective [11]. MIP can tirelessly apply social pressure to large groups, and thus create a more 'intense' form of persuasion.

Though it has been established that normative social influence can be used effectively in persuasive technology [12], it is unknown if the application of more social pressure also makes it more effective. The aim of this research is to study whether more intense persuasion is also more effective when utilizing normative social influence in persuasive technology.

The organisation of this paper is as follows. In section two we focus on related work on social influence and persuasive technology, in section three on the different roles social influence can take in persuasive technology, and in section four on a definition of effectiveness and intensity. These sections establish the foundations for the experiment that is described in the remainder of this paper, starting with the method for the experiment in section five. The results of this experiment are presented in section six, followed by the discussion of these results in section seven.

2 Related Work on Social Influence and Persuasive Technology

When persuasion within groups is mentioned, conforming rather than compliance is implied. This means that a person is more susceptible to be persuaded when he or she has positive feelings towards the persuaders. But even when this person would be persuaded by a group of strangers, the pressure to conform is strong. So strong, that some set aside their own judgement and conform to the judgement of the majority, even if the judgement of the majority is clearly wrong, as was demonstrated by the Asch conformity experiment [13].

This can further be explained by the social impact theory by Latané [14]. He suggested that the impact of social pressure is determined by three factors: strength, immediacy and number. Here, strength is the power, importance or intensity of the source of the pressure to the target. Immediacy being the closeness of the source to the target, in both space and time. And number meaning the number of people pressuring the target. According to this theory, you feel more social pressure from a large group of relatives in the same room than from a small group of strangers that pressure

you by sending a letter. Although Bond suggests that the impact of the number factor decreases for larger groups [15].

The social impact theory and the principles discussed in this section describe how social influence can be used for persuading people to show certain behaviour. In addition to this, research by Fogg has shown that to persuade a person to change his or her behaviour using technology, the person has to be sufficiently motivated, have the ability to perform the behaviour, and has to be triggered to perform the behaviour [16]. For a person to be persuaded and show the desired target behaviour, these three factors — motivation, ability and trigger — have to be present at the same time.

Of the three factors, the motivation factor can be split into three types of motivators: pleasure or pain, hope or fear, and social acceptance or rejection. Each pair represents both a positive and a negative motivation. Which type of motivator can be used best depends on many factors, including the type of technology used and the type of target behaviour that a person needs to show. The focus of this research lies on the social acceptance or rejection motivator, on which the social influence principles described earlier in this section are applicable.

3 Comparing and Mediating Technology

As mentioned in the previous section, three factors — motivation, ability and trigger — have to be present at the same time in order for a person to be persuaded using technology. We argue that, in addition to the three pairs of motivators described earlier, there are also two main roles technology can take to trigger someone. We say that technology takes a *comparing* role if it makes a person aware of the behaviour of others, or that it takes a *mediating* role if it makes the person aware that others are aware of his or her behaviour.

Consider the following example. Scott often forgets to call his grandmother. It's not that Scott does not want to call his grandmother. On the contrary, he loves his family very much. He just does not think about it, unlike Scott's relatives, who call grandmother at least weekly. Recently, Scott's family remarked that he doesn't call his grandmother very often. Because of this, Scott feels pressed to give his grandma a call. He doesn't want to lose his image of a good grandson and therefore promises to call her more often.

To trigger Scott to call more often taking a comparing role, we would show him how often he calls his grandmother in comparison to his friends and family. For example, we could put the number of 'granny calls' behind the names of friends and family in the address book of his mobile phone. Whenever Scott makes a call, he would be made aware of his behaviour in comparison to others.

If we take a mediating role, we show Scott's friends and family how often he calls his grandmother. Not only does this raise the pressure on Scott because of the way other people may think of him ("I should definitely give grandma a call before my family starts to think bad of me."), but others may also pressure Scott in real life to call his grandmother more often ("Scott, I see that you haven't called grandma for weeks, so...").

In both cases there are several ways to raise or lower the pressure on Scott or, in other words, to vary the intensity of the persuasion. This research is focussed on the comparing role.

4 Defining Effectiveness and Intensity

To be able to answer the question if more intense social influence is also more effective when using persuasive technology, the effectiveness and the intensity of persuasion have to be defined. Effectiveness can be measured in two ways: by measuring the increase in target behaviour of individual people, or by measuring the percentage of people that show the target behaviour. It should be noted that both ways of measuring effectiveness may also be used at the same time, by measuring the increase in the number of times people show certain target behaviour, while at the same time measuring the percentage of people that show this increase in behaviour.

There are many ways to persuade someone to show a certain target behaviour using normative social influence. For example, when persuading someone to come along to the movies, one could ask "Would you like to join me to the movies?", but one may also ask "Would you like to join us to the movies? All our friends are coming too." The normative social pressure present in the second question makes the pressure to conform bigger, since all your friends are joining as well. Therefore, we argue that the second attempt to persuade is more intense.

The intensity of persuasion is dependent of a large number of variables. Ranging from the emotional state of the person that persuades [17] to the speed at which this person speaks [18]. Therefore, this research focuses on those variables that are linked to normative social influence and that can be controlled with persuasive technology. All other variables are kept as constant as possible. This led to the following variables, mainly based on the social impact theory by Latané.

1. Strength (or peer relationship). The closeness of the people in a group to the target of the persuasion. The value of this variable is based on the theory by Granovetter [19] and can be strong tie ('friends, family'), weak tie ('acquaintances'), or no tie ('strangers').
2. Number (or group size). The total number of people in a group.
3. People pressuring. The number of people in a group that are persuading the target. The value of this variable could either be the minority or the majority of the people in the group.
4. Message frequency. The number of times one or more people in the group pressure the target of the persuasion. Cacioppo and Petty established that repeating a message could have a positive influence on the effect of an attempt to persuade [20].

In this paper it is hypothesized that a closer relationship, a larger group, a higher number of people pressuring, or more frequent messages means more intense persuasion. A more distant relationship, a smaller group, a lower number of people pressuring, or less frequent messages means less intense persuasion.

The main hypothesis for this research is therefore: "When utilizing normative social influence in persuasive technology, more intense persuasion is more effective". To falsify this hypothesis, the experimental results were analysed with the following null hypotheses:

1. "Participants of the groups that were persuaded ($C_{1\text{-}4}$) and the control group (C_0) are equally likely to continue". Rejecting this null hypothesis would indicate that the persuasion in general had an effect.
2. "Participants that were persuaded subject to the conditions 'minority' (C_1, $_3$) and 'majority' (C_2, $_4$) are equally likely to continue". Rejecting this null hypothesis would indicate that the *people pressuring* variable has an effect.
3. "Participants that were persuaded subject to the 'ties' conditions (C_3, $_4$) and 'no ties' conditions ($C_{1,\ 2}$) are equally likely to continue". Rejecting this hypothesis would indicate that the *strength* variable has an effect.
4. "Participants that were persuaded subject to any of the conditions $C_{1\text{-}4}$ or any other of the conditions $C_{1\text{-}4}$ are equally likely to have answered the most questions after normative social influence". Rejecting the null hypothesis would indicate that suggesting a norm of five or more questions has an effect.

5 Method

To test the main hypothesis, a quantitative experiment on the online social network Facebook was conducted. The main goal of the experiment was to see whether people could be persuaded to continue to answer more questions after they first decided to stop, by using various levels (i.e. intensities) of normative social influence.

5.1 Participants

Participants were recruited by sending a Facebook or e-mail message with an invitation to participate in an experiment, falsely described as studying how photographs on Facebook influence emotions. Participants were required to have a Facebook account, and could choose to take part in either a Dutch or English version of the experiment. The authors of this paper used their personal networks to recruit participants.

5.2 Design

Setting up an experiment that uses both normative social influence and persuasive technology is quite complex. It should be flexible enough to vary the amount of normative social influence while at the same time remain believable for the participants. It was decided to use an online social network, Facebook, as the setting for the experiment. It allowed the use of existing social networks, allowing easy data collection and manipulation of variables.

It was also decided not to disclose the real intentions of the experiment to reduce the risk of participants recognizing the persuasion as the main goal of the experiment. Instead, participants were told they participated in a study on the effects that

photographs have on emotions. This façade-experiment, as it will be called here, consisted of showing a series of 200 photographs depicting scenes from nature, people and architecture. Participants were asked which emotions they felt while looking at the photographs. They could select two emotions from two drop-down lists containing emotions, ranging from positive to negative. For example, a possible answer could be: "This picture makes me feel sad and disgusted". The answers to these questions were however not relevant for the actual experiment.

It was emphasized that the participants could rate as few or as many photographs as they chose to (although it was actually limited to 200 photographs) and that they could press the 'stop and send' button whenever they wanted to stop and submit their results. This way the participants had no norm for the number of photographs to be rated. When the 'stop and send' button was pressed, a message was displayed that attempted to persuade the participant to answer more questions using normative social influence, i.e. to conform to the norm supposedly set by other participants (comparing role, see section 3).

The application for the experiment was built with the scripting language PHP in combination with the Facebook API and photographs collected using Google Images[1].

A participant was randomly subjected to one of five conditions C_0 – C_4 below. The content of the message was dependant on the condition:

- C_0 (control group): The participant is not persuaded by means of normative social influence. When a participant decided to stop, the following message was displayed: "Would you like to answer some more questions or do you really want to stop?"
- C_1: Normative social influence by displaying the following message when a participant decided to stop: "Are you sure you want to stop? 11% of the participants in this experiment have answered 5 more questions than you. Would you like to answer some more questions or do you really want to stop?"
- C_2: Normative social influence by displaying the following message when a participant decided to stop: "Are you sure you want to stop? 87% of the participants in this experiment have answered 5 more questions than you. Would you like to answer some more questions or do you really want to stop?"
- C_3: Normative social influence by displaying the following message when a participant decided to stop: "Are you sure you want to stop? 11% of your Facebook friends who participated in this experiment have answered 5 more questions than you. Would you like to answer some more questions or do you really want to stop?"
- C_4: Normative social influence by displaying the following message when a participant decided to stop: "Are you sure you want to stop? 87% of your Facebook friends who participated in this experiment have answered 5 more questions than you. Would you like to answer some more questions or do you really want to stop?"

[1] Used options: image size: 800x600, image type: Photo, usage rights: Only images labeled for reuse.

These conditions were based on four levels of intensity that are shown in Table 1, which in turn were based on the four variables described in Section 4: the peer relationship, group size, people pressuring and the message frequency. For this experiment the message frequency variable was normalised to one message and the group size variable was not used.

Table 1. The four levels of intensity of normative social influence that were used in this experiment. The control group (C_0) was not influenced.

	Facebook friends	**Other participants**
11% of people *(minority)*	11% of Facebook friends (C_3)	11% of other participants (C_1)
87% of people *(majority)*	87% of Facebook friends (C_4)	87% of other participants (C_2)

Majority and minority were expressed using the percentages of 87% and 11%, respectively. These percentages were made up and are not based on actual data from other participants (and therefore do not add up to 100%). Peer relationships were expressed as 'Facebook friends' (ties) and 'other participants' (no ties).

The target behaviour for the persuasion in this experiment was: *Increasing the number of questions a participant answered*. The comparative trigger was formed by the persuasive message that appeared once when the participant indicated a desire to stop answering questions. The assumed motivation of the participants was the social acceptation of answering just as many, or more, questions as the other participants. The required ability to answer more questions was the press of a button.

To measure the effectiveness of this persuasion the following was measured: the number of people that decided to answer more questions (after indicating a desire to stop), the number of questions that were answered, and the time spent before and after persuasion.

5.3 Procedure

The participants first had to log in with their Facebook account through a secure connection. After logging in, participants were asked to judge a series of photographs.

When a participant pressed the 'stop and send' button, one of the messages was presented to verify the action (see the conditions in the design section for the messages). After a time-out of 5 seconds, to prevent participants from proceeding without reading the message, participants could either choose to actually stop and end the façade-experiment or continue to answer questions. When a participant chose to continue, the experiment continued as before, until the participant would press the 'stop and send' button again. At that moment no message was shown and the experiment ended.

Upon finishing the façade-experiment, participants were thanked for their participation and were asked to remark on things they noticed during the façade-experiment in a form field. One month after commencement, the façade-experiment was ended and each participant was informed about the real intentions of the experiment.

The experiment ran from December 7th, 2010 until January 6th, 2011.

6 Results

The 98 Facebook users[2] were classified as 39 females, 41 males and 18 persons with an unknown gender[3]. The age ranged from 19 – 61 years, with an average of 27 years ($sd \approx 7$). Of the participants, 79 had a Dutch and 19 an English language preference. Almost the same percentage of Dutch and English speaking participants continued to answer more questions (respectively 31.7% and 31.6%). A T-test could not detect a significant difference in mean number of questions answered between Dutch and English participants ($t = 0.36$; $df = 96$; $p = 0.72$). Table 2 shows an overview of the general data that were registered during the experiment.

Table 2. General data registered during the experiment

	N	Range	M	SD
Age	91	19 – 61	27.43	6.67
Questions answered before initial stop	98	1 – 159	31.69	27.05
Time spent before initial stop *(in min.)*	98	0.25 – 38.93	8.96	7.95
Questions answered after initial stop	31	0 – 59	13.77	12.91
Time spent after initial stop *(in min.)*	31	0.08 – 12.57	2.84	2.90

Is there a difference in the number of participants that continued between the control group (C_0) and the groups that were persuaded ($C_{1\text{-}4}$)?

A Mann-Whitney U-Test rejected the null hypothesis 1 ($p = 0.039$). We therefore conclude that in the experimental data there is a significant difference between using no persuasion (C_0, $n = 18$) and persuasion ($C_{1\text{-}4}$ combined, $n = 80$).

Is there a difference in the number of participants that were persuaded to continue between the 'minority' conditions ($C_{1,3}$) and 'majority' conditions ($C_{2,4}$)?

Although on average slightly more participants continued with the 'majority' condition (Table 3), the performed Mann-Whitney U-Test did not reject the null hypothesis

[2] 10 of the 108 participants were excluded: 4 participants logged in but answered no questions; 3 participants did not finish the experiment; 3 participants remarked that they became aware of the real intentions of the experiment.

[3] The gender was automatically read from the Facebook profiles, but not all profiles contained this information.

2 ($p = 0.23$). We may therefore not conclude from the experimental data that the number of peers has a significant influence on the effectiveness of the persuasion.

Is there a difference in the number of participants that were persuaded to continue between the 'ties' conditions ($C_{3,4}$) and 'no ties' conditions ($C_{1,2}$)?

Although on average slightly more participants continued subject to the 'ties' than to the 'no ties' conditions (Table 4), the performed Mann-Whitney U-Test did not reject the null hypothesis 3 ($p = 0.39$). We may therefore not conclude from the experimental data that the relationship between the participant and the (pretended) persuader has a significant effect on the effectiveness.

Do the groups conform to the norm of answering five more questions?

Although there are differences in mean number of questions answered before and after various levels of normative social influence (Table 3) were exercised (in particular for the number of questions after the message for condition C_2 ($m \approx 18$, $sd \approx 13$), a Mann-Whitney U-Test did not reject the null hypothesis 4 for any of the condition pairs (excluding C_0). The least likely null hypothesis (i.e. the one closest to rejection, $p = 0.11$) is for the comparison of the groups subject to the least intense persuasion C_1 and most intense persuasion C_4. Nonetheless, we may not conclude from the experimental data that there is a significant effect of the intensity of persuasion on the number of questions answered.

Table 3. An overview of the continuation and the mean number of questions answered before and after the message

Condition	N	Mean no. of questions before message	No. of participants that continued	Mean no. of questions after message
C_0 *(control)*	18	25.50 *(sd 23.23)*	2 *(11.1%)*	29.50 *(sd 41.72[4])*
C_1 *(min., no tie)*	20	28.70 *(sd 21.41)*	7 *(35.0%)*	7.86 *(sd 6.39)*
C_2 *(maj., no tie)*	19	33.27 *(sd 31.92)*	9 *(47.4%)*	17.67 *(sd 13.11)*
C_3 *(min., tie)*	23	30.70 *(sd 17.89)*	6 *(26.1%)*	11.17 *(sd 9.95)*
C_4 *(maj. tie)*	18	40.72 *(sd 38.59)*	7 *(38.9%)*	12.43 *(sd 6.05)*
Total	98	31.69 *(sd 27.05)*	31 *(31.6%)*	13.77 *(sd 12.91)*

[4] The control group has a very high standard deviation because one participant answered 59 questions while another participant answered no questions at all. It therefore has no importance.

7 Discussion

The results of the experiment indicate that normative social influence can be successfully applied to persuasive technology, which supports our hypothesis and earlier findings by Fogg [10].

We could not find significant differences in effectiveness between a minority and a majority of peers ($C_{1,2}$ and $C_{3,4}$) on the α 5% level. However, although it is not significant, the majority conditions ($C_{2,4}$) do seem to be more effective than the minority conditions ($C_{1,3}$). Furthermore, our results do not allow us to conclude that the peer relationship has influence on the effectiveness of persuasion.

Although we came fairly close to rejecting the null hypothesis comparing between the least intense and most intense persuasion ($C_{1,4}$), it cannot be concluded that when utilizing normative social influence in persuasive technology, more intense persuasion is more effective. The effect, if any, does not seem to be strong enough to, with our limited number of participants, detect a significant influence of the intensity levels that were used on the number of questions answered.

These results are most easily explained by the limited number of participants, and the high variability of their responses. The high and non-Gaussian variability also led us to use non-parametric Mann Whitney U-tests and a matching relatively broad null hypothesis. This makes the null hypothesis harder to reject and if rejected, gives less strong evidence to our main hypothesis, but the tests require far less assumptions and are much more robust to outliers.

During the experiment we also found some weaknesses in its design. Aspects like the way sentences are constructed or how usable the software is from an HCI (human-computer interaction) perspective have influence on the results.

The percentages used to show minorities and majorities might also have had undesired effects, as one participant mentioned that when she was confronted with the message that 11% of her Facebook friends answered more questions, she thought that she was actually doing quite well since most of her Facebook friends answered *less* questions then her. The same goes for the message "87% of the participants in this experiment have answered 5 more questions than you", as participants might have thought that these other participants would (at least partially) include their friends[5].

The results may also have been influenced by a lack of believability of the façade-experiment. Especially for those participants who were aware that they share few or no Facebook friends with other participants. On the other hand there were several indications from the received feedback that the participants *did* believe the façade-experiment. We received multiple feedbacks along the lines of: "The possible choices to answer my feelings do not reflect the way I feel". Though this qualitative data is of course difficult to objectively quantify.

The experiment described in this paper should be repeated in different ways and with different conditions and above all with more participants, in order to see if the intensity of persuasion *does* have influence on effectiveness. Judging from the results

[5] It should also be noted that Facebook friends do not necessarily have to be real friends. They may also be relatives or total strangers.

of the experiment, this will be a worthwhile effort that could shine new light on the ways normative social influence can be used in persuasive technology, and how this might lead to more effective ways to help people change their behaviour, feelings or thoughts.

Acknowledgements. The authors of this paper would like to thank Rogier Brussee for his extensive help with the statistical analysis of our research data and with editing the final draft of this paper, Joost Broekens, Arjan Haring, and Maurits Kaptein for their support and advice, B.J. Fogg and Jason Hreha for sharing their paper on target behaviours, Sanne Duhoux, Krystina Poludnikiewicz and Niels Reijmer for proofreading the final paper, and all of the participants of the experiment.

References

1. Remmerswaal, J.: Handboek groepsdynamica: een nieuwe inleiding op theorie en praktijk. H. Nelissen, Soest, The Netherlands (2003)
2. Hogg, M., Vaughan, G.: Social Psychology. Pearson Education Limited, Harlow (2008)
3. Deutsch, M., Gerard, H.B.: A study of normative and informational social influences upon individual judgement. Journal of Personality and Social Psychology 51, 629–636 (1955)
4. Zimbardo, P.G., Leippe, M.: Psychology of Attitude Change and Social Influence. McGraw-Hill, New York (1991)
5. Fogg, B.J.: Persuasive technology: using computers to change what we think and do. Morgan Kaufman Publishers, San Fransisco (2002)
6. Fogg, B.J.: Persuasive computers: perspectives and research directions. In: Proc. of the SIGCHI Conf. on Human Computer Interaction, pp. 225–232 (1998)
7. Winterboer, A., Cramer, H.S., Pavlin, G., Groen, F.C., Evers, V.: Do you smell rotten eggs? In: Proc. of the 11th Int. Conf. on Human-Computer Interaction with Mobile Devices and Services, MobileHCI 2009, pp. 1–2 (2009)
8. Bărbat, B., Zamfirescu, C.B., Costache, G.: Agent-Oriented Captology For Medical Informatics. In: Proc. of the MIE 2000, pp. 465–469 (2000)
9. Lucero, A., Zuloaga, R., Mota, S., Muñoz, F.: Persuasive Technologies in Education: Improving Motivation to Read and Write for Children. In: IJsselsteijn, W.A., de Kort, Y.A.W., Midden, C., Eggen, B., van den Hoven, E. (eds.) PERSUASIVE 2006. LNCS, vol. 3962, pp. 142–153. Springer, Heidelberg (2006)
10. Fogg, B.J.: Mass Interpersonal Persuasion: An Early View of a New Phenomenon. In: Oinas-Kukkonen, H., Hasle, P., Harjumaa, M., Segerståhl, K., Øhrstrøm, P. (eds.) PERSUASIVE 2008. LNCS, vol. 5033, pp. 23–34. Springer, Heidelberg (2008)
11. Ploderer, B., Howard, S., Thomas, P., Reitberger, W.: Hey World, Take a Look at Me!: Appreciating the Human Body on Social Network Sites. In: Oinas-Kukkonen, H., Hasle, P., Harjumaa, M., Segerståhl, K., Øhrstrøm, P. (eds.) PERSUASIVE 2008. LNCS, vol. 5033, pp. 245–248. Springer, Heidelberg (2008)
12. Michael Weiksner, G., Fogg, B.J., Liu, X.: Six Patterns for Persuasion in Online Social Networks. In: Oinas-Kukkonen, H., Hasle, P., Harjumaa, M., Segerståhl, K., Øhrstrøm, P. (eds.) PERSUASIVE 2008. LNCS, vol. 5033, pp. 151–163. Springer, Heidelberg (2008)
13. Asch, S.E.: Opinions and social pressure. Scientific American 193, 31–35 (1955)
14. Latané, B.: The psychology of social impact. American Psychologist 36, 343–356 (1981)

15. Bond, R.: Group size and conformity. Group Processes and Intergroup Relations 8, 331–354 (2005)
16. Fogg, B.J.: A behavior model for persuasive design. In: Proc. of the 4th Int. Conf. on Persuasive Technology, Persuasive 2009, pp. 1–7 (2009)
17. Wood, W.: Attitude change: persuasion and social influence. Annual Review of Psychology 51, 539–570 (2000)
18. Miller, N., Maruyama, G., Beaber, R.J., Valone, K.: Speed of speech and persuasion. Journal of Personality and Social Psychology 34, 615–624 (1976)
19. Granovetter, M.: The Strength of Weak Ties: A Network Theory Revisited. Sociological Theory, 105–130 (1982)
20. Cacioppo, J., Petty, R.: Effects of message repetition on argument processing, recall, and persuasion. Basic and Applied Social Psychology 10, 3–12 (1989)

Exploring Perceived Persuasiveness of a Behavior Change Support System: A Structural Model

Filip Drozd[1,2], Tuomas Lehto[3], and Harri Oinas-Kukkonen[3]

[1] Changetech AS, Research & Development, Gaustadalléen 21, N-0349 Oslo, Norway
[2] Department of Psychology, University of Oslo, P.O. Box 1094 Blindern, N-0317 Oslo, Norway
[3] Oulu Advanced Research on Software and Information Systems (OASIS), Department of Information Processing Science, University of Oulu, 90570 Oulu, Finland
fd@changetech.no,
{tuomas.lehto,harri.oinas-kukkonen}@oulu.fi

Abstract. There is no healthcare system in the world that has the capacity or resources to provide every person in need of help and support of changing lifestyle behaviors. Consequently, there is a need to design health information systems that enable individuals to manage their health and maintain a healthier lifestyle. However, there is limited knowledge about how individuals perceive these behavior change support systems and how individuals' perceptions affect the use of such systems. In the present study, we tested a persuasive systems design model that had a significant impact on perceived persuasiveness and system usage. Also, there appears to be some local gender differences in the strength of the relationships between factors (perceived persuasiveness and intention, and unobtrusiveness and intention). We discuss future developments of the model and health as a social and personal responsibility.

Keywords: persuasive systems design, behavior change support systems, usage, gender and technology, eating habits, weight loss, partial least squares.

1 Introduction

According to the World Health Organization's [1] projections for 2030, about 51.5 million deaths or 76 percent of the global mortality and burden of disease will be accounted for by non-communicable conditions (i.e. diseases which are largely caused by poor lifestyle and health behaviors). There is no healthcare system in the world that has the capacity or resources to provide every individual in need of help and support of changing lifestyle and health behaviors. New innovations are therefore very much needed and consumer health applications can potentially help to this end. However, many kinds of information system (IS) have been developed and designed primarily for healthcare managers and professionals. Thus, there is a need for health information systems that enable individuals to manage their health and maintain a healthier lifestyle. There is an increasing interest in reaching consumers and patients directly through consumer health IT. According to Payton and colleagues [2] (p. vi),

M. Bang and E.L. Ragnemalm (Eds.): PERSUASIVE 2012, LNCS 7284, pp. 157–168, 2012.

there has been "*a shift in the role of the patient from passive recipient to active consumer of health information and active user of healthcare devices, logging, and monitoring systems*". Indeed, by providing consumers with access and tools to personal health information, we can begin to influence how they manage their health and well-being.

Oinas-Kukkonen [3] proposed the generic concept of a behavior change support systems (BCSS) to describe consumer health applications. BCSSs highlight autogenous and voluntary approaches in which people use information technologies to change their attitudes or behaviors through building upon their personal motivation or goal. They harness either technology-mediated persuasion or technology-human persuasion. Technology-human persuasion is fully automatized, whereas technology-mediated persuasion means that people are influencing others through e.g. discussion forums, instant messages, or social network systems. The primary challenge in developing such comprehensive systems for consumers, is that there is modest knowledge of how individuals interact with consumer health informatics and how they process and act on information [4] (i.e. how individuals perceive these systems and how individuals' perceptions affect the use of such systems). In a recent report by Jimison and colleagues [5], the most frequent barrier to use of interactive consumer health IT across studies, was a lack of perceived benefit and lack of convenience. Furthermore, subjects were less likely to use systems if they did not fit seamlessly into their regular daily routines. Other major obstacles were burdensome data entries and lack of trust in the provided information. Clearly, technologies cannot have the capacity to facilitate self-monitoring and self-management or improve consumers' health outcomes if consumers do not accept the technology e.g. [6].

The objective of this study is to investigate consumers' perceptions of a web-based intervention for weight loss. Specifically, we aim to examine factors affecting the perceived persuasiveness of the system and whether perceived persuasiveness predicts intention to use the intervention and actual system usage. First, a theory-driven research model is constructed. Second, a component-based structural equation modeling (SEM) approach, partial least squares (PLS), is used to test the relationships between a latent variable and its indicators (i.e. the measurement model) and the structural relationships among the latent variables in the research model.

2 Theoretical Background and Research Model

The interaction between people and IT is an area of inquiry that accentuates the multi-disciplinary nature of the IS field. Human behavior impacts the whole life cycle of IT; including its design, development, deployment, adoption, and use. In this study, we were interested in a web-based intervention designed to influence its users' eating behaviors and built a theory-driven research model based on the Persuasive Systems Design[1] (PSD) model (see Figure 1) [7], to test the interventions perceived persuasiveness and how perceived persuasiveness of the intervention relates to usage.

[1] The "social support" category from the PSD model [9] has been omitted from the proffered research model, since the web-based program under investigation does not facilitate social support in its current form.

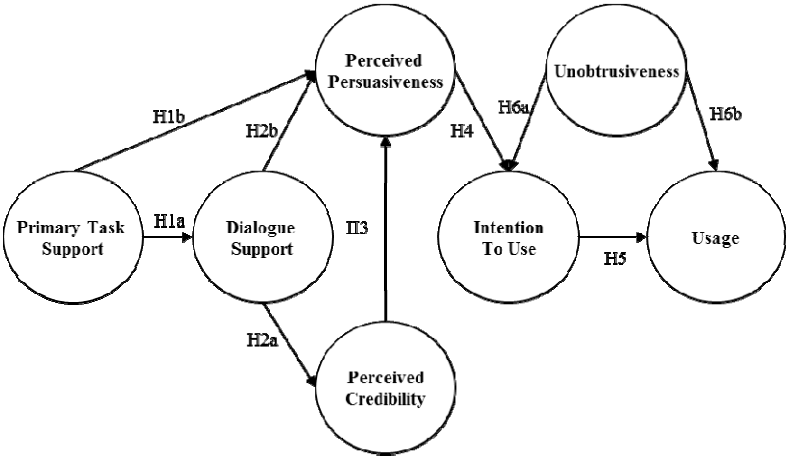


Fig. 1. The PSD model with hypothesized relationships between constructs in the present study

Primary Task Support. Primary task support encompasses the means to aid the individual in performing his or her primary task [7]. The aim of primary task support is to enhance the self-efficacy of the user and to reduce the cognitive burden and disorientation towards system use [10], [9]. According to Johnston and Warkentin [10] (p. 3), self-efficacy is "*the degree to which an individual believes in his or her ability to enact the recommended response*." In addition, primary task support increases positive affect [11] which augments the persuasiveness of the source [12]. Therefore, we put forward the following hypotheses:

H1a: Primary task support is positively related to dialogue support.
H1b: Primary task support positively affects perceived persuasiveness.

Dialogue Support. People tend to react to IT applications as if they were interacting in social situations [13], [14], [15]. Thus, supporting the dialogue between the IT application and the individual user is essential. Dialogue support defines the key principles in keeping the user active and motivated in using the system, and involved in his or her behavior change process. System-to-user prompts, praise and reminders play an important role in dialogue support [7]. The dialogue support may be further enhanced, for instance, by providing users with appropriate counseling and feedback. Such instances of dialogue support promote users' engagement, motivation, and positive affect, which will likely influence users' confidence in the source (credibility). Consequently, the following hypotheses are formulated:

H2a: Dialogue support positively influences perceived credibility.
H2b: Dialogue support positively affects perceived persuasiveness.

Perceived Credibility. Perceived credibility contains both a subjective and objective component. The subjective component is based upon people's initial evaluations of the system credibility based on their first impression. An encounter with a (novel) system is largely a visual one, and during system interaction, constant visual

information immediately elicits aesthetic judgments. This principle is called surface credibility [7]. Perceived credibility also has an objective component which might be bolstered by providing endorsements from respected and renowned sources (e.g. a recommendation by an authoritative organization, an award for excellence in usability, or a privacy seal to ensure confidentiality). Both the subjective and objective component of perceived credibility should affect the believability of the IT application. Consequently, we state the following hypothesis:

H3: Perceived credibility positively affects perceived persuasiveness.

Perceived Persuasiveness. In classical models of attitude change, messages are presented, received, processed, and if successful, recipients' attitudes shift toward the advocated position [16]. The altered attitude may have an impact on subsequent behavior under appropriate conditions [16] but, according to Crano and Prislin [16], a central aspect that must be taken into account when reflecting on persuasion involves the fundamental construct of attitude. They state the following (p. 347): "*Today, most accept the view that an attitude represents an evaluative integration of cognitions and affects experienced in relation to an object.*" In the present study, perceived persuasiveness is operationally defined as individual's favorable impressions toward the system. The following hypothesis is rendered:

H4: Perceived persuasiveness has a positive impact on intention to use the system.

Intention to Use and Usage. Behavioral intentions are proposed by several psychological theories, as an immediate and important predictor of behavior. In IS research, behavioral intentions are often used as a proxy for system usage or net benefits of an IT system. The problem of using only behavioral intentions as an outcome is that intention is not always a good predictor of behavior. Meta-analyses show that intentions do have a significant impact on behaviors but typically explain just under 30 percent of the variation across different types of behaviors [17], [18]. Therefore, it was important to include the link from intention to actual system usage in this study:

H5: Intention to use the system at two weeks predicts actual usage after six weeks.

Unobtrusiveness. To understand and close the intention-behavior gap, much of the attention of research has been drawn to various perspectives such as self-regulation [19] or intervention characteristics [20]. According to the PSD model, unobtrusiveness may be one such important factor that bridges some of the intention-behavior gap. Technology provides the information and means to aid the individual in his or her tasks but the key to successful implementation and use, depends on whether users have the opportunity to use the system or whether they find it disturbing. According to Oinas-Kukkonen and Harjumaa's PSD model [7], systems should aim at unobtrusiveness. Unobtrusiveness is a contextual construct that reflects whether the system fits with the user's environment in which he or she uses the system. On the one hand, research shows how important it is to have a fit between technology and its

users on individual performance [21]. On the other hand, intrusive technology characteristics are found to have negative consequences such as stress [22]. Consequently, we hypothesized that:

H6a: Unobtrusiveness has a positive relationship to intention to use the system.
H6b: Unobtrusiveness has a positive relationship to actual usage of the system.

3 Research Method

3.1 Data Collection and Subject Characteristics

Subjects were recruited through online ads and banners over a period of 14 days during October 2011. By clicking on a banner, potential subjects were redirected to an external website containing study information and an informed consent. Subjects had to confirm to have read the study information before they could proceed to fill in the online survey. Subjects with a verified e-mail address, ≥ 18 years, and < 5 missing values, were included in the dataset. Data were collected online at baseline, two weeks, and six weeks post-intervention. The surveys consisted of questions related to 1) demographics, 2) primary task support, 3) dialogue support, 4) perceived credibility, 5) perceived persuasiveness, 6) unobtrusiveness, and 7) intention to use. A seven-point Likert scale was applied for all continuous items (ranging from strongly disagree to strongly agree). Usage was collected by means of log-file data six weeks post-intervention, about the time as users would have finished the program with optimal program compliance. Overall, 128 complete responses were obtained. See Table 1 for detailed information regarding the respondents.

Table 1. Baseline sample characteristics

Characteristic	Females (n=64)	Males (n=64)	Total (n=128)
Age (yrs)	37.6 ± 12.3	41.8 ± 10.5	39.7 ± 11.6
Education			
Elementary	1 (1.6)	2 (3.1)	3 (2.3)
High-school	16 (25.0)	17 (26.6)	33 (25.8)
1-3 yrs college or university	16 (25.0)	20 (31.3)	36 (28.1)
4-5 yrs college or university	15 (23.4)	12 (18.8)	27 (21.1)
> 5 yrs college or university	16 (25.0)	13 (20.3)	29 (22.7)
Occupational status			
Employed	36 (56.3)	56 (87.5)	92 (71.9)
Unemployed	1 (1.6)	0 (0.0)	1 (0.8)
Student	16 (25.0)	4 (6.3)	20 (15.6)

Note. Numbers represent mean and ± SD for age and number of observations with percentage of observations in parenthesis for education and occupational status.

3.2 Description of the Behavior Change Support System

Ned i Vekt is a fully automated web-based behavior change support system developed by Changetech AS. The aim of the program is threefold: 1) assist users changing their

eating habits, 2) up-regulate positive emotions and mood, and 3) losing weight. It is a tunneled program consisting of two program days for six weeks. Every Monday and Thursday, the user receives an e-mail with a link to the day's website. As shown in table 2, every day in the program is unique and consists of psycho-educative information, online exercises, and home assignments. Much of the program content is based on consumer psychology [23], positive psychology [24], and the basic premises of the non-dieting paradigm [25], i.e., 1) stable mild and moderate overweight is not unhealthy, 2) dieting is ineffective, and 3) dieting is harmful.

Table 2. Overview of program days in Ned i Vekt

Day	Psychoeducative information	Exercise(s)
1	Food and emotions	Test of eating behaviors; personal reasons for changing eating behaviors
2	Willpower (focus on one thing at the time)	Implementation intentions and optimism exercise
3	Eating environment and "forbidden" foods	Suppression-countersuppression experiment (ironic mental processing)
4	Willpower, blood sugar levels, and performance	Savoring positive moments
5	Eating environment	How environmental factors such as lighting, temperature, music, distractors, etc., affect our eating behaviors
6	Temptations and impulses	Attentional control ("cold spots" exercise) and stereotype lift
7	Food and expectations	Demonstration of the size-contrast and vertical-horizontal illusions
8	Associations with food (e.g. affect heuristics)	Relaxation training
9	Self-efficacy and change	Positive self-talk and mindful eating (the raisin exercise)
10	Stress, willpower, and choice of foods	Exemplar priming (i.e. a story with word primes for increased performance)
11	Summary and repetition	Test of eating behaviors

4 Data Analysis and Results

We analyzed our research model using partial least squares (PLS) by utilizing WarpPLS 2.0 (Scriptwarp Systems; www.scriptwarp.com/warppls/) software for data analysis. WarpPLS is a component-based path modeling software application which is appropriate to use when the purpose of the model is to predict, rather than to test established theory [26]. Moreover, PLS is reasonably robust to deviations from a multivariate distribution [27]. The statistical objective of PLS is similar as that of linear regression, i.e., to demonstrate explained variance in the latent variable as indicated by R^2 values, to indicate the strength of the relationship between latent variables in terms of β-values, and test the significance of the relationship between latent variables by estimating t-values and reporting their corresponding p-value [27]. It is often suggested that the minimal sample size in PLS analysis should be at least 10 times the number of indicators in the most complex construct. Our total sample

size exceeded this requirement. However, we applied the jackknifing procedure to generate more stable re-sample path coefficients. Overall, testing the PLS model is carried out in two steps: 1) the assessment of the reliability and validity of the measurement model, and 2) the assessment of the structural model. The measurement model includes the relationships between the constructs and the indicators used to measure them. The convergent and discriminant validity of the research instrument is examined in order to verify that the constructs' measures are valid and reliable before attempting to draw conclusions regarding relationships among constructs (structural model). The structural model includes testing the full research model in a single step.

The Measurement Model. Descriptive statistics for the research constructs are presented in Tables 3 and 4. The properties of the scales are assessed in terms of item loadings, discriminant validity, and internal consistency. Item loadings and internal consistencies greater than .70 are considered acceptable [28]. The values presented in Table 3 have been obtained through IBM SPSS Statistics 19 software. All constructs in the model display good internal consistency as evidenced by their composite reliability scores (from .90 to .97) and Cronbach's alpha scores (from .84 to .96).

Table 3. Construct means and reliability scores for total sample (n=128)

Construct	No. of items	Mean ± SD	Composite reliability	Cronbach's alpha
Primary task support (PRIM)	3	14.7 ± 4.0	.94	.91
Dialogue support (DIAL)	4	19.3 ± 5.4	.95	.93
Perceived credibility (PCRED)	5	26.7 ± 5.5	.95	.94
Perceived persuasiveness (PERS)	4	20.3 ± 5.3	.93	.84
Unobtrusiveness (UNO)	4	21.2 ± 5.0	.90	.91
Intention to use (INTE)	4	23.8 ± 5.7	.97	.96
Usage (USE)	1	10 (2–11)	.94	

Note. Usage was measured with an ordinal single indicator based on how many program days a user had completed. Thus, usage is reported as median with range in parenthesis.

Table 4. Latent variable (LV) correlations for total sample (n=128).

LV	AVE	PRIM	DIAL	PERS	PCRED	UNO	INTE	USE
PRIM	.84	.92						
DIAL	.83	.84	.91					
PERS	.78	.78	.83	.88				
PCRED	.80	.69	.69	.77	.90			
UNO	.69	.48	.57	.61	.51	.83		
INTE	.90	.56	.60	.67	.51	.60	.95	
USE	1.0	.24	.32	.32	.22	.31	.31	1.0

Notes. The principal diagonal (shaded cells) is the square root of the AVE (Average Variance Extracted) between the constructs and their measures. Off-diagonal figures are the inter-construct correlations. For discriminant validity, the principal diagonal should be greater than off-diagonal elements.

The Structural Model. The total sample size was 128; however, in order to examine the female and male sub-groups individually, we employed jackknifing. Kock [29] suggests that for small samples (i.e. < 100 subjects), jackknifing is the recommended re-sampling approach. For consistency, we also tested the total sample with jackknifing procedure. Figure 2 provides the results of the PLS analysis for the final model in the total sample. For the most part, there is substantial support for the model, although the relationships between primary task support and perceived persuasiveness (β = .13, p = .10), and unobtrusiveness and usage (β = .18, p = .18) were not supported. All of the significant hypotheses were supported at p < .001. Primary task support accounts for 72 percent of the variance in dialogue support (β = .85), while dialogue support accounts for 48 percent of the variance in perceived credibility (β = .69). Dialogue support (β = .55) and perceived credibility (β = .39) together explain a substantial amount, 76 percent, of the variance in perceived persuasiveness. Fifty-five percent of the variance in intention to use program is explained by perceived persuasiveness (β = .52) and unobtrusiveness (β = .31). Finally, intention to use program (β = .32) explains actual program usage but accounts for only 10 percent of the variance.

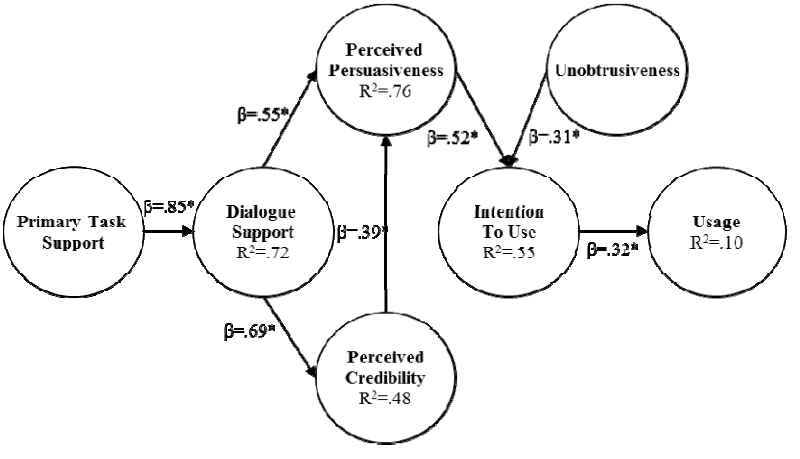


Fig. 1. A path diagram for the total sample (n=128) displaying the final results of the PLS analysis for the PSD model (*p<.001)

Further analyses between female and male sub-groups identified the same relationships between constructs as for the total sample. There were slightly different R^2 values among females and males (see table 5). The largest R^2 difference between the sub-groups occurred in the relationship between primary task support and dialogue support (PRIM→DIAL). For the female subgroup the R^2 is 68 percent, whereas for the male sub-group it is 76 percent, indicating that primary task support explains more variance in dialogue support for males. However, most interesting, are the findings that the relationship between perceived persuasiveness and intention (PERS→INTE) was stronger for males (β = .55) than for females (β = .37), and that the relationship between unobtrusiveness and intention (UNO→INTE) was stronger for females (β = .45) than males (β = .32).

Table 5. Paths, coefficients, and R^2 values for the final models

	β-values (weights)			R^2		
Path(s)	Females	Males	Total	Females	Males	Total
PRIM→DIAL	.83**	.87**	0.85**	.68	.76	.72
DIAL→PCRED	.68**	.72**	0.69**	.46	.52	.48
DIAL→PERS	.53**	.59**	0.55**			
PCRED→PERS	.42**	.35**	0.39**	.76	.78	.76
PERS→INTE	.37*	.55**	0.52**			
UNO→INTE	.45**	.32**	0.31**	*.57*	*.58*	*.55*
INTE→USE	.35*	.43*	0.32**	.12	.19	.10

Notes. Jackknifing was used for re-sampling. *p < .05, **p < .01.

It would seem that the R^2 difference on intention was attenuated because of these inverted relationships between perceived persuasiveness and intention, and unobtrusiveness and intention, among males and females. Overall, the differences between R^2 values appear to be rather modest between males and females, however, the path coefficients suggests moderating effects of gender on the relationships between perceived persuasiveness and intention, and unobtrusiveness and intention. Thus, the data indicate slightly different patterns of how the information system is perceived among males and females. This may explain why intention accounts for more variance in usage among males (19%) than females (12%).

5 Discussion

The persuasive systems design categories in the PSD model [7] appear to have a significant impact on perceived persuasiveness and actual system usage. The results supported most of the hypothesized relationships between factors that affect the perceived persuasiveness and system usage, except the direct relationship between primary task support and perceived persuasiveness. The results also demonstrate that contextual factors such as unobtrusiveness, directly affect intention but not actual usage. Furthermore, the final PSD model was replicated across female and male sub-groups at the global level. Admittedly, there appears to be local gender differences in perceptions of the system, most notably, in the strength of relationships between perceived persuasiveness and intention, and unobtrusiveness and intention. Significant gender differences have previously been documented in perceptions of website design and website satisfaction [30] and online trust [31]. More such differences between genders in perceptions and usage of IT can be expected, especially since some of these differences appear to be biologically founded, see [32].

We argue that dialogue support (system-to-user and user-to-system) is a crucial factor for the persuasiveness of IT systems and acts as a connecting node to other related factors. Dialogue support has three major connections: primary task support, perceived credibility and perceived persuasiveness. Through dialogue support, users receive appropriate feedback and counseling, which keeps them motivated, engaged, and involved in their change process. Low dialogue support would not only appear to

result in low motivation to behavior change, but to have a negative impact on the perceived persuasiveness of the entire system. As Lyytinen [33] noted, computers are no longer merely reactive devices that are optimized to respond to user requests but more proactive. Current technological advances allow novel solutions for dialogue support, such as embodied conversational agents [11] or even persuasive robotic assistants [34].

This study has a number of limitations. First, this study represents an exploratory test of a theoretical model and should be subject to further testing with various participants, technologies, and contexts. Second, the relationship between primary task support and other factors in the PSD model are not yet fully clear and they need to be validated more thoroughly. Third, research subjects were from one country, so the results may not generalize to other settings and contexts.

6 Conclusions

This paper tested a theory-based model predicting factors contributing to perceived persuasiveness and actual usage of a consumer health application. Researchers and designers in e-health may benefit from this type of approach to promote IT adoption and usage. Clearly, the enormous costs in healthcare demand for innovative solutions for various stakeholders in healthcare [see 2]. By providing consumers with access to personal health information, we can begin to influence and support self-management of health. From a societal point of view, people's health is not only a social responsibility; it is also a personal responsibility which affects other people and the available capacity and resources in the healthcare system. From an academic perspective, results of this research will contribute to the IS and e-health literatures by developing an IT adoption model for persuasive behavior change support systems. It is hoped that this research will attract the attention of researchers to further develop and test constructs and models applicable to consumers' use of health information systems as preventive health measures. From a more practical viewpoint, we argue that studying the adoption, use and impact of novel consumer health IT is feasible as it will guide future implementations.

Acknowledgments. We thank Prof. Pål Kraft for his input to the study design. TL and HOK would like to thank SalWe Research Program for Mind and Body (Tekes – the Finnish Funding Agency for Technology and Innovation grant 1104/10).

Conflicts of Interest. The first author is working for Changetech AS which has developed the web-based program under investigation.

References

1. Organization, W.H.: Projections of mortality and burden of disease (2004-2030), http://www.who.int/healthinfo/global_burden_disease/projections/en/index.html
2. Payton, F.C., Pare, G., Le Rouge, C.M., Reddy, M.: Health care IT: Process, people, patients and interdisciplinary considerations. Journal of the Association for Information Systems 12, i–xiii (2011)

3. Oinas-Kukkonen, H.: Behavior Change Support Systems: A Research Model and Agenda. In: Ploug, T., Hasle, P., Oinas-Kukkonen, H. (eds.) PERSUASIVE 2010. LNCS, vol. 6137, pp. 4–14. Springer, Heidelberg (2010)
4. Eysenbach, G.: Consumer health informatics. British Medical Journal 320, 1713–1716 (2000)
5. Jimison, H., Gorman, P., Woods, S., Nygren, P., Walker, M., Norris, S., Hersh, W.: Barriers and drivers of health information technology use for the elderly, chronically ill, and underserved. AHRQ Publication No. 09-E004, pp. 1–1422. Agency for Healthcare Research and Quality, Rockville, MD (2008)
6. Or, C.K.L., Karsh, B.-T.: A systematic review of patient acceptance of consumer health information technology. Journal of the American Medical Informatics Association 16, 550–560 (2009)
7. Oinas-kukkonen, H., Harjumaa, M.: Persuasive systems design: Key issues, process model, and system features. Communications of the Association for Information Systems 24, 485–500 (2009)
8. Nadkarni, S., Gupta, R.: A task-based model of perceived website complexity. MIS Quarterly 31, 501–524 (2007)
9. Webster, J., Ahuja, J.S.: Enhancing the design of web navigation systems: The influence of user disorientation on engagement and performance. MIS Quarterly 30, 661–678 (2006)
10. Johnston, A.C., Warkentin, M.: Fear appeals and information security behaviors: An empirical study. MIS Quarterly 34, 549–566 (2010)
11. Derrick, D.C., Jenkins, J.L., Nunamaker Jr., J.F.: Design principles for special purpose, embodied, conversational intelligence with environmental sensors (SPECIES) agents. AIS Transactions on Human-Computer Interaction 3, 62–81 (2011)
12. Angst, C.M., Agarwal, R.: Adoption of electronic health records in the presence of privacy concerns: The elaboration likelihood model and individual persuasion. MIS Quarterly 33, 339–366 (2009)
13. Nass, C., Steuer, J., Tauber, E.R.: Computers are social actors. In: Proceedings of the Conference on Human Factors in Computing Systems - CHI 1994, pp. 72–78. ACM Press, Boston (1994)
14. Al-Natour, S., Benbasat, I.: The adoption and use of IT artifacts: A new interaction-centric model for the study of user- artifact relationships. Journal of the Association for Information Systems 10, 661–685 (2006)
15. Lee, E.-J.: I like you, but I won't listen to you: Effects of rationality on affective and behavioral responses to computers that flatter. International Journal of Human-Computer Studies 67, 628–663 (2009)
16. Crano, W.D., Prislin, R.: Attitudes and persuasion. Annual Review of Psychology 57, 345–374 (2006)
17. Armitage, C.J., Conner, M.: Efficacy of the Theory of Planned Behaviour: A meta-analytic review. British Journal of Social Psychology 40, 471–499 (2001)
18. Sheeran, P.: Intention-behavior relations: A conceptual and empirical review. European Review of Social Psychology 12, 1–36 (2002)
19. de Ridder, D.T.D., de Wit, J.B.F.: Self-regulation in health behavior: Concepts, theories, and central issues. In: de Ridder, D.T.D., de Wit, J.B.F. (eds.) Self-Regulation in Health Behavior, pp. 1–24. Wiley, Chichester (2006)
20. Webb, T.L., Sheeran, P.: Does changing behavioral intentions engender behavior change? A meta-analysis of the experimental evidence. Psychological Bulletin 132, 249–268 (2006)

21. Goodhue, D.L., Thompson, R.L.: Task-technology fit and individual performance. MIS Quarterly 19, 213–236 (1995)
22. Ayyagari, R., Grover, V., Purvis, R.: Technostress: Technological antecedents and implications. MIS Quarterly 35, 831–858 (2011)
23. Wansink, B.: Mindless eating: Why we eat more than we think. Bantam, New York (2006)
24. Lyubomirsky, S.: The how of happiness: A scientific approach to getting the life you want. Penguin, New York (2007)
25. Polivy, J., Herman, C.P.: Undieting: A program to help people stop dieting. International Journal of Eating Disorders 11, 261–268 (1992)
26. Chin, W.W., Marcolin, B.L., Newsted, P.R.: A partial least squares latent variable modeling approach for measuring interaction effects: Results from a Monte Carlo simulation study and an electronic-mail emotion/adoption study. Information Systems Research 14, 189–217 (2003)
27. Gefen, D., Straub, D.W., Boudreau, M.-C.: Structural equation modeling and regression: Guidelines for research practice. Communications of the Association for Information Systems 4, Article 7 (2000)
28. Fornell, C., Larcker, D.F.: Evaluating structural equation models with unobservable variables and measurement error. Journal of Marketing Research 27, 39–50 (1981)
29. Kock, N.: Using WarpPLS in e-collaboration studies: Mediating effects, control and second order variables, and algorithm choices. International Journal of e-Collaboration 7, 1–13 (2011)
30. Cyr, D., Bonanni, C.: Gender and website design in e-business. International Journal of Electronic Business 3, 565–582 (2005)
31. Sanchez-Franco, M.J., Villarejo Ramos, A.F., Martin Velicia, F.A.: The moderating effect of gender on relationship quality and loyalty toward Internet service providers. Information & Management 46, 196–202 (2009)
32. Riedl, R., Hubert, M., Kenning, P.: Are there neural gender differences in online trust? An fMRI study on the perceived trustworthiness of eBay offers. MIS Quarterly 34, 397–428 (2010)
33. Lyytinen, K.: HCI research: Future directions that matter. AIS Transactions on Human-Computer Interaction 2, 22–25 (2010)
34. Loije, R., Neerincx, M.A., Cnossen, F.: Persuasive robotic assistant for health self-management of older adults: Design and evaluation of social behaviors. International Journal of Human-Computer Studies 68, 386–397 (2010)

Biometric Monitoring as a Persuasive Technology: Ensuring Patients Visit Health Centers in India's Slums

Nupur Bhatnagar[1,⋆], Abhishek Sinha[1], Navkar Samdaria[2], Aakar Gupta[2], Shelly Batra[1,*], Manish Bhardwaj[3], and William Thies[4,*]

[1] Operation ASHA
[2] University of Toronto
[3] Innovators In Health
[4] Microsoft Research India
{nupur.bhatnagar,shelly.batra}@opasha.org, thies@microsoft.com

Abstract. Managing chronic disease is particularly challenging in the developing world, because every trip to a health center can translate to lost time and wages on the part of the patient. This problem is especially acute for tuberculosis patients, who in India are required to visit a center over 40 times in the course of a six-month treatment period. In this paper, we explore the role of a biometric attendance terminal in persuading patients to complete follow-up health visits in slum communities of New Delhi, India. The terminal, which enrolled over 2,300 patients across 25 centers during our 2 years of observation, uses biometric fingerprint scanning to ensure that tuberculosis patients receive and take medications on the right schedule. We evaluate the perceived impact of the terminal via interviews with 8 health workers, 4 center owners, and 23 patients. Our findings suggest that the biometric terminal helps to draw patients to the center, both by incentivizing health workers to convince patients to come, and by persuading patients that in-person visits are important.

1 Introduction

Persuading patients to comply with prescribed treatment can often make the difference between life and death. This is especially true for tuberculosis, a disease that remains the largest infectious killer of adults, claiming nearly 1.8 million lives annually [5]. In India, which is responsible for nearly one fifth of the global TB burden [5], these deaths are especially tragic because there are free and effective antibiotics available from the government. However, in order to be cured, an infected patient needs to take the drugs (initially 7 pills at a time) on a strict schedule: three days per week continuously over a 6- to 8-month period. Failure to complete this regimen can lead to drug resistance, which is typically deadly – not only for the patient, but also for others he or she might infect.

In order to persuade patients to adhere to tuberculosis treatment regimens, India and other countries have adopted a program known as Directly Observed Therapy, Short Course (DOTS). In this program, patients ingest each dose of medication under direct observation. The observers, known as medication *providers*, keep the medications with

⋆ Corresponding authors.

M. Bang and E.L. Ragnemalm (Eds.): PERSUASIVE 2012, LNCS 7284, pp. 169–180, 2012.

them in order to closely control their distribution and administration. Patients are therefore required to travel to a DOTS center to receive each dose. While the DOTS program has had a significant impact on tuberculosis control in India, it also introduces a new challenge, which is to persuade patients to make regular visits to the center. Due to myriad factors spanning education, travel, work schedules, side effects, and forgetfulness, it remains difficult to convince patients to visit the center regularly, thereby jeopardizing their treatment outcome.

To address this problem, we designed and deployed a biometric attendance terminal that uses fingerprinting technology to record every visit that a patient makes to a health center. Every day, the visitation log is uploaded via SMS to a central location, where program managers can analyze the data and offer targeted counseling or supervision for patients who have missed their dose. We deployed this terminal across 25 tuberculosis treatment centers in slum communities in urban India. To date, our deployment spans 2 years and encompasses over 2,300 patients. We evaluated the impact of the terminal via interviews with health workers, center owners, and patients.

Our results indicate that health workers perceive the biometrics platform as improving patient attendance at the health center, thereby improving their medication adherence and prospects for recovering from TB. One of our key findings is regarding the impact of biometrics on the social dynamics between health workers and patients. While we expected the terminal to strain the trust of this relationship, due to the vigilant monitoring introduced, our anecdotal evidence suggests the opposite. In their conversations with patients, health workers describe the biometric platform as a common adversary to which they are also bound, thereby creating empathy in the patient and persuading them to make regular visits out of solidarity with the health worker. While some health workers and patients acknowledge that biometrics can cause inconvenience, and not all patients are able to explain the purpose of the system, these drawbacks are minor compared to the benefits cited. Overall, our findings suggest that biometric attendance tracking can play a positive role in promoting healthy behaviors in low-income settings.

2 Related Work

Ensuring that patients adhere to prescribed medication is a pressing global problem that has been studied extensively [9]. A review of interventions seeking to improve adherence found 17 programs that had statistically significant effects on treatment outcomes [6]. Successful interventions were usually complex, incorporating more convenient care, information, counseling, reminders, self-monitoring, reinforcement, family therapy, and other forms of additional supervision or attention. However, none of these studies addressed adherence to TB medications in a developing-country context.

Even under directly-observed treatment regimens, it is recognized that many factors prevent adherence to tuberculosis medications [2]. A study focusing specifically on the Indian context found that men dropped out of treatment because the visits jeopardized work and earning opportunities, while women dropped out due to housework and the strain of keeping their condition secret, particularly when women's movement outside of the home is questioned [8]. Prior researchers have worked to improve DOTS therapy using videophones in place of in-person visits [3], a strategy that is judged to save costs

in the United States but is still prohibitively expensive in the context of slum communities in India. Researchers have also worked to reduce the need for DOTS by using medication monitors to log self-administered doses [7], though this approach remains controversial amongst policymakers due to the difficulty of monitoring treatment and also the potential resell of government-funded TB medications. Prior work in persuasive technology has also explored the design of an electronic medication monitor [13].

Biometric technologies are receiving increasing attention as a means to verify the delivery of social services to underprivileged populations. This dialogue is especially relevant in India, where the UID (Unique Identification) project aims to provide a unique biometric identity to each of the 1.2 billion citizens of the country. Proponents of UID point to numerous benefits for poor communities, including increased access to bank accounts, decreased waste due to undeserved welfare payouts, and, recently, increased access to healthcare. The UID authority goes so far as to say that "health related development schemes could offer a killer application for the UID" [14]. However, this claim is not without controversy. Critics from academia, the media, and the public sphere have been vocal in questioning the role of biometrics in healthcare delivery. One researcher concludes that "the UID scheme has thus little to offer for improvement in the public health situation in the country" [11].

While the debate regarding the role of biometrics in healthcare delivery is becoming increasingly relevant, and increasingly polarized, there is surprisingly little scientific evidence regarding the impact of biometrics on real-world health programs. Research on biometric identification in the developing world has focused on health clinics in South Africa [12], anti-retroviral therapy in Malawi [16], clinical trials in Vietnam [4], and tracking of nomadic pastoralists in Chad [15]. However, to the best of our knowledge, there is no prior study that examines the interplay of biometrics and healthcare delivery for low-income populations in India. This paper seeks to fill this gap by examining a large-scale deployment of biometric technologies in health clinics in slum communities.

3 Biometric Attendance Tracking

The context for our project is a tuberculosis treatment program run by Operation ASHA (OpASHA), one of the premier TB treatment organizations in the world. OpASHA serves over 8,000 patients across 11 Indian states, including a longtime focus on slum communities in Delhi. They utilize a highly innovative model that includes all the elements of government DOTS programs, as well as some additions. In an OpASHA DOTS center, there are two actors: a *provider* and a *counselor*. The provider is a member of the local community who offers physical space for use as a DOTS center throughout the day. The counselor is a full-time OpASHA employee who administers medication for several hours per day from the provider's location, and also delivers doses to patients' homes when they fail to come to the center. Since patients visit centers only 3 days per week, counselors typically are responsible for two centers, which they visit on alternate days. While OpASHA's model has been highly successful, they sought to improve the transparency of DOTS center operations to enable them to scale across India and other countries.

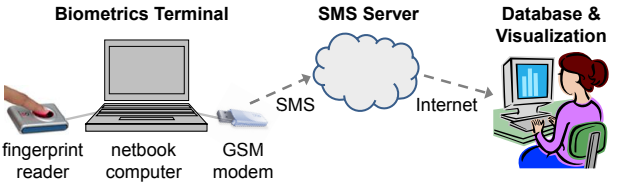


Fig. 1. Overview of our biometric attendance system

3.1 Technology Design

Our approach in this project is to utilize a biometric fingerprinting terminal to provide a transparent and tamperproof record of which doses of TB medication are administered to a patient under supervision of a health worker. While similar records are recorded on paper today, they are widely regarded as being unreliable and unrepresentative of daily operations; also, they are difficult to promptly aggregate and analyze. Using a biometric terminal with daily uploads to a central location, it becomes possible for program managers to identify missed doses as soon as they occur, and to respond with improved counseling or supervision.

The technical details of our system have been described in a prior publication [10]. Briefly, the system consists of three components: a netbook computer, a biometric fingerprint reader, and a GSM modem (see Figure 1). When patients visit the center, they scan their fingerprint, and the system logs their visit. At the end of the day, the history of visits is uploaded via SMS to a server, where it becomes available for visualization and analysis by program managers.

3.2 Technology Deployment

The biometric terminal was deployed to all 17 of Operation ASHA's centers in Delhi and, starting in October 2011, to 8 centers in Jaipur (see Figure 2). In this process, the terminal was taken up by 13 counselors and enrolled over 2,300 patients. Since there is a separate terminal for each counselor and provider, we deployed a total of 38 terminals.

Deploying a new technology at this scale in a resource-poor environment is no small feat. Training the counselors and attaining their commitment to the system was the first hurdle. During training, we dispelled fears, such as the counselors being replaced by the computers or being micro-monitored by the management. As counselors had limited exposure to computers, they thought they might damage the system and were hesitant to use it. We needed to hold multiple training sessions, and offered on-site help for initial patient enrollment, before counselors were comfortable using the system on their own.

The second hurdle that emerged over time was the recognition accuracy of the fingerprint reader. Though we used a state-of-the-art device (the Digital Persona U.are. U4500), our conditions were very demanding, as our low-income clientele often had calloused fingers, and we were also using the reader to identify them without any assistance (i.e., the system *identified* patients based on their fingerprint, rather than *verifying* their fingerprint relative to another form of identification.) Though mis-recognitions

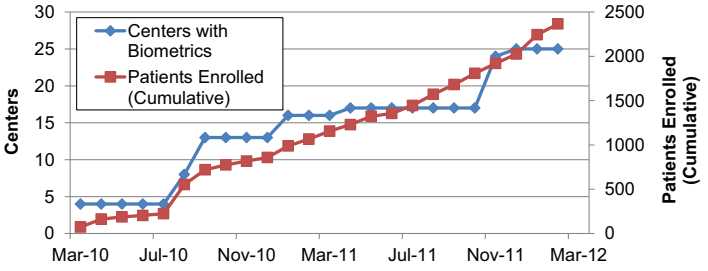


Fig. 2. Deployment timeline of the biometrics terminal

were relatively rare, we aimed to eliminate them entirely. We accomplished this by requiring two fingerprint scans per patient. If the system matches the print to the same person on both trials, then they are logged in as that person. Otherwise, the patient is asked to try again until two successive scans are consistent.

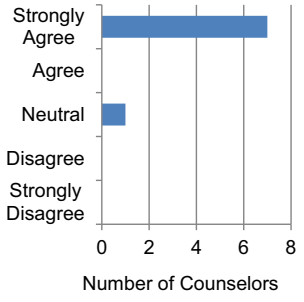
The final challenge in our deployment was dealing with various software and hardware failures of the laptops, spanning computer viruses, power surges, physical damage, and bugs in our initial platform. Of these, computer viruses proved to be the biggest challenge. Despite its intended use as a healthcare platform, we learned the laptops were frequently used for entertainment purposes during off hours, introducing viruses via USB thumb drives. While antivirus software proved insufficient to curb the issue, we eventually stabilized the systems by using Windows Steady State to prohibit most types of changes from being stored to disk.

4 Perceived Impact of Biometrics

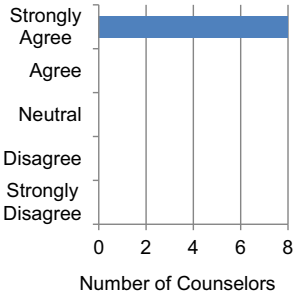
Our goal in this study was to evaluate the potential of biometrics to improve health outcomes in a tuberculosis treatment program. While the ideal methodology for such a study would be to conduct a randomized controlled trial and measure the impact of biometrics on missed doses, default rate, and cure rate, such a trial is beyond the scope of our initial investigation. While we did examine these outcomes across the centers studied, the sample size was too small, and there were too many confounding factors, to identify any statistically significant effect of the intervention. While we look forward to conducting a more rigorous trial in the future, in this work we focus on qualitative feedback as an indicator of the potential impact of the biometric system.

4.1 Methodology

To understand the perceptions of the biometrics technology by various stakeholders in the treatment program, we conducted semi-structured interviews with 8 counselors, 4 providers, and 23 patients. Interviewers were conducted by the first author, a technology officer of Operation ASHA, who had led the deployment of biometrics over the previous 15 months and was already known to the counselors and providers. While the interviewer made every effort to remain impartial, some participants' responses may have been biased by a desire to please the interviewer. We revisit this issue in Section 5.



(a) "Patients are more likely to visit a center which has biometrics."



(b) "I am less likely to send medication home with a proxy (other than the patient) due to the biometric terminal."

Fig. 3. Counselors' agreement with interview statements

Interviews were conducted over two months (Jul-Aug 2011) and took place in either the DOTS center or the Operation ASHA office. Interviews were done in Hindi and typically lasted between 10 and 45 minutes. The audio was recorded and later transcribed and translated to English. The authors then analyzed the (anonymized) English transcripts for common themes and patterns, which are summarized below.

4.2 Perceptions of Counselors

We interviewed eight counselors, spanning all of Operation ASHA's active centers in Delhi. Counselors were predominantly female (one was male), had an average age of 27, and had an average household income of Rs. 14,500 ($295) per month. Counselors were relatively well-educated, with all completing 12th standard and three holding a B.A. Most counselors (5 out of 8) had some training with computers, usually via a 3 to 12 month course. However, only a fraction of counselors (3 out of 8) had any real-world experience with computers. All counselors owned a mobile phone, as well as a TV. Counselors were employed with Operation ASHA for an average of 1.6 years (min 6 months, max 3.3 years) at the time of our interviews. They had been using biometrics in their centers for an average of 7 months (min 4 months, max 1.1 years).

Biometrics helped to persuade patients to come to the center. Of all the points cited in favor of biometrics, counselors were most vocal about the benefits in drawing patients to the centers. Ensuring regular meetings between patients and counselors is the cornerstone of the DOTS treatment protocol. However, these meetings often break down. In the words of one counselor, "Many patients want unsupervised doses and when I refuse they even offer me money. They don't want to come to the center, but send someone else in their place."

As illustrated in Figures 3a and 3b, counselors are almost unanimous in feeling that the biometrics caused more patients to visit the centers, and also reduced the likelihood of sending medication home with a different person (the 'proxy'). One counselor explains this benefit as follows:

> Earlier someone else would come with some excuse and pick the medicine. We could not argue too much because the patients would threaten to quit the treatment or worse

> skip medication. Now even they realize that it's a rule from the government and they can't avoid it. The patients have become more active and we have also become more alert and vigilant.

We gathered numerous anecdotes in which counselors believed that certain patients would not have visited the centers if it weren't for the biometric devices. For example:

> N. was pulmonary positive and was also very sick. He lived very far off and was not ready to come to the center. I went to his house for first counseling session and registered him in the laptop. I motivated him, giving the reason of biometrics. So for almost a month till he was too sick to walk, he came with an attendant on a rickshaw. He was excited to see biometrics and kept coming in regularly to give his fingerprint. He was fit to move about within 2 months.

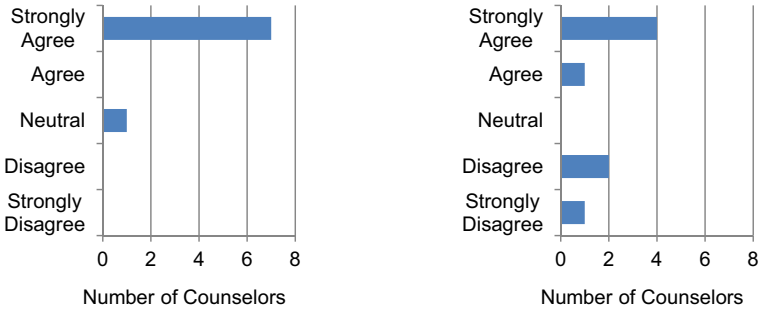
Even when patients are in better health, obligations such as work can prevent them from visiting the centers. The biometric devices have also been effective in drawing such patients in. As noted by one counselor, "A patient goes on motorcycle to Noida for work [consuming his entire day] and wanted someone else to collect his doses. When I showed him the fingerprint device and told him that it would not be possible, he took half day off on 3 days of the week."

In other cases, it is the counselor who takes extra measures to ensure that medication is delivered. For example, one counselor took regular trips to a patient's residence to ensure that their interaction was logged: "There is a handicapped patient who is unable to come to the center. And he wants unsupervised doses. Since I can't give it to a proxy but still I want to give him DOTS, I go myself every time and take his fingerprint."

The biometrics also had an interesting and unexpected impact on counselor-patient interactions. Previously, counselors alone were responsible for enforcing the DOTS protocol, and if patients disliked the protocol, they directed their negative feelings at the counselors. However, upon introducing the biometrics, it became clear to patients that both counselors and themselves were subject to monitoring by a higher authority (the central office), thereby reducing tensions between counselors and patients. In the event of a dispute, counselors can 'blame' the biometrics for enforcement of the protocol, while retaining good relations with the patient. In the words of one counselor, "All patients come to the DOTS center, some out of consideration for me as I have told them that I get scolded if scans are not taken." Another counselor confirms, "If a patient doesn't want to come and gives the excuse of going for duty, I tell them that this is important and that this record goes to the office. This becomes a reason for them to visit the center."

Counselors viewed biometrics as improving the treatment program. Overall, the counselors were in broad agreement that the biometric attendance terminals were improving the main activities of the treatment program, namely 1) enforcing adherence to DOTS, as patients were more likely to visit the centers themselves instead of sending a 'proxy', and 2) helping counselors to identify and follow-up with patients who had not yet visited on a given day. Together, these benefits also imply that counselors had more interaction with patients, further improving care. Figure 4a illustrates that 7 out of 8 counselors felt strongly about the overall improvement to the program.

One of the key benefits is ensuring that patients retrieve their own medication, as summarized by one counselor: "It has benefited us. Our interaction with the patient



(a) "Biometrics is improving the treatment program."

(b) "Biometrics enables me to do my job faster."

Fig. 4. Counselors' agreement with interview statements

increases because the patient is forced to come to the center." Another counselor emphasizes the book-keeping benefits of biometrics: "I would be lost without biometrics. I would not know how many patients to expect that day, how many and who have already visited the center."

The counselor responsible for the neutral reply in Figure 4a was frustrated by some technical difficulties in registering and recognizing patients on the terminal. However, this same individual "strongly agreed" that medication was less likely to be sent by proxy when using the biometrics.

Counselors earned more respect using biometrics. In addition to benefiting the treatment program, the biometric devices also offered auxiliary benefits to counselors. As summarized by one counselor, "I have started getting more respect from all" since the introduction of the laptops. Another counselor explained, "Now that I have this laptop the patients give me double respect. When I go into the field even the neighbors of the patient flock around and think I am coming from a big hospital because I carry a laptop. So they listen carefully and even ask questions."

Two counselors noted increased respect within their families, as well. One counselor said, "In my family, now my value has increased. My hubby did not like that I was working among TB patients. They feel good now that I have been given this responsibility [of a laptop]."

While these perceptions do not directly impact patients, they are important for the uptake of biometric devices within the organization. The success of the biometric program ultimately rests with the counselors, who are the main mediators of the technology. As it often requires patience and perseverance for them to learn how to use the technology, the respect gained can provide an important incentive and reward for carrying through with the program.

Biometrics increased the amount of work for some counselors. While the feedback from counselors was largely positive, our discussion would not be complete without considering the possible drawbacks of biometrics. One area that could be improved further is the efficiency of counselors with the technology. As illustrated in Figure 4b,

counselors had mixed opinions as to whether the biometric devices enabled them to do their jobs more quickly.

There are several explanations for this result. First, as the biometric technology is in a pilot stage, we are currently requiring counselors to maintain *both* paper and electronic records, so that the paper records can serve as a backup in the event of any unexpected behavior from the biometrics software. As stated by one counselor, "Biometrics has doubled my work as I have to maintain both the records now."

Another potential challenge to the acceptance and sustainability of the biometrics is that it holds counselors accountable for very strict adherence to the treatment protocol. When asked if she would be happy if the biometric system is withdrawn, one counselor replied, "I will be happy because I am under a lot of tension that if patients are not given unsupervised doses, they will quit and come in default and if I give unsupervised then they will come in the missed dose list." While adherence to the protocol is critical for successful implementation of DOTS, counselors who were previously unaccustomed to monitoring could find it more stressful to be subject to constant auditing via biometrics. We expect that this feeling may prove to be different amongst new counselors, who did not have prior experience in an unmonitored treatment scenario.

4.3 Perceptions of Providers

We interviewed four providers. Providers were male, had an average age of 36 (min 30, max 43) and had an average household income of Rs. 17,500 ($389) per month. All providers had a health diploma. Half had some background with computers, while half had some informal knowledge; none had received formal training. All providers owned mobile phones and TVs. Providers had worked with Operation ASHA for an average of 1.5 years (min 6 months, max 4 years) and had been using biometrics in their centers for an average of 11 months (min 5 months, max 1.25 years).

Providers believed that patients were more likely to visit due to biometrics. Like the counselors, providers perceived that the presence of biometrics increased the likelihood that patients would visit their DOTS center. Three out of four providers strongly agreed with this idea, while one was neutral. One provider said, "I think it a very good initiative. This has helped us convince more patients to come to the DOTS center. Initially the counselor went to some patients' homes to take the fingerprints. Now the patients understand that it is important that the records in the laptop are up-to-date. So they come without much ado."

Because more patients visit the DOTS center during the daytime, there are fewer after-hours visits for the provider to handle. One provider said, "The counselor makes sure that she gets most of the fingerprints when she is on duty and there are fewer patients that are left pending for the day."

Providers preferred to have the biometrics in their centers. Overall, providers were strong advocates of the biometrics program. Three out of four providers strongly agreed that they prefer to keep the technology in their centers, while one was neutral. It is not obvious that providers would have this sentiment, since they are held liable for a laptop that is kept on the premises at all times. They were responsible for keeping it safe from theft and tampering.

4.4 Perceptions of Patients

We interviewed 23 patients (11 male), who were members of slum communities. Seven were illiterate, while the rest had completed an average of 9 years of schooling. The majority of female patients (9 of 12) were housewives. Other patients were students (n=4) or employed in small jobs such as selling vegetables or helping in an office. The average household income amongst patients was Rs. 8300 ($184) per month. Almost all of the patients (21 out of 23) owned a mobile phone, and most (17 out of 23) owned a TV. They were enrolled with OpASHA's treatment program for an average of 3.9 months (min 2 months, max 8 months).

Patients had varied understandings of the purpose and benefit of biometrics. While the intent of the biometric system was explained briefly to patients by the counselor, the purpose and benefits of biometrics were not understood in all cases. This may have been compounded by the patients' unfamiliarity with biometric technology. Only two patients had used biometrics before (one for attendance in an office, and one for registering for an insurance scheme).

We asked patients to explain why the biometrics were in use. Sixty-one percent of respondents understood well, for example, saying that "it is for my benefit so that I come every day and give my fingerprint." However, 39% of respondents could not articulate the intent of the system. About half of patients (52%) expressed positive feelings for the biometrics, saying that it was "good" or similar, while the remaining patients did not offer an opinion.

When asked if the biometrics was helping the treatment, some patients perceived an impact. For example: "It is definitely helping the patients. Because of laptops the patients are coming to the center, and when they are here they will take their medicine also." However, others were not clear on how the system helps the patient. Several said that they "can't say" whether the biometrics helps the patients, while one patient, perhaps in an attempt to please the interviewer, said "Yes [it helps], but I can't say how."

When asked if the biometrics helped them personally, one patient said that her primary motivation in visiting the DOTS center was not due to biometrics, but due to better health: "I don't know [if the system helps me]. I would have come even if this device wasn't there because I want to get well." At the same time, another patient admitted that the biometrics changed her behavior in visiting the DOTS center: "I think so. If laptop wouldn't have been here I may not have come to the center so regularly but would have sent my husband."

Several patients were impressed that the organization was investing expensive resources in their treatment, and in turn felt more motivated to visit centers. A counselor recounted a story in which one patient encouraged another by saying, "The government and organization is doing so much to treat us – free medicines, free tests, even facility of attendance system to ensure that we get well. In spite of this if we can't come to the center to treat ourselves then we don't deserve to be treated!"

At the same time, some patients acknowledged that due to the biometric monitoring, they were under increased pressure to visit the DOTS centers, and this could sometimes lead to inconvenience. One counselor recalled, "The drivers have a problem coming to the center as they have erratic hours of duty. Earlier they could come on non-schedule day and take their medicine. Now they are under compulsion to come themselves and

on the scheduled day. Such patients don't like this system." We view this sentiment as evidence that the biometrics is having its intended effect. While the treatment protocol may be inconvenient at times, when followed closely it ensures that all patients are cured.

5 Study Limitations

One limitation of our study is that responses of participants may have been biased due to *demand characteristics*, which is the tendency for participants to adjust their responses to conform with the investigator's expectations. Since our interviewer was a known champion of the biometric terminals, and also held an influential position in the organization, respondents could have felt pressure to speak positively about the technology.

To account for this contingency, we performed a second round of (shorter) interviews in which a different interviewer posed one of the same questions to the counselors. This interviewer (the second author) also contributed to the biometrics project, but had less interaction with the counselors to date and was thus less of an authority figure in their eyes. He asked counselors, "would you be disappointed if biometrics was removed from the program?" While 7 counselors responded similarly to before, one counselor reversed her stance, confessing that she was "a little scared" to admit to the first interviewer that she didn't want the terminal. She said the terminal was heavy and difficult to carry. This experience prompted us to study demand characteristics in more detail, culminating in an independent demonstration of dramatic response bias when working with underprivileged populations in India [1].

Unfortunately it is too late for us to precisely characterize or compensate for the influence of response bias on the results reported in this paper. We believe that most of the sentiments are genuine, especially those supported by concrete anecdotes. However, an independent assessment using interviewers that are not associated with the project, and not of greater social status than the counselors, could help to corroborate our findings.

6 Conclusions

This paper examines the potential benefits and hazards of deploying a biometric attendance terminal in a real-world healthcare setting. Unlike prior studies, our work focuses on low-income communities in India, where the role of biometrics is currently an active subject of scrutiny and debate. While our findings are preliminary, and based only on the opinions of stakeholders in the program, they do suggest that biometrics can play a positive role in persuading patients to adhere to a treatment protocol – in this case, ensuring regular consumption of tuberculosis medications.

Our interviews suggest two primary mechanisms by which biometrics is affecting behaviors. First, it persuades the counselor: by offering a transparent and tamperproof record of attendance, it makes counselors more accountable and thus more incentivized to attract patients to the center. Second, it persuades the patient: as most patients understand that the biometrics system cannot be fooled, they come to the center to avoid causing any negative repercussions for the counselor, or for themselves. There may also be a third element of persuasion, that we have yet to study, which is the influence on program managers as they start to incorporate biometric reports into their daily supervision of staff.

Future research could further investigate the relative role and interplay of these mechanisms. However, our goal in this paper was not to isolate the mechanisms of persuasion, but rather to understand whether biometrics, on the whole, can make a real tuberculosis treatment program more likely to succeed. In the future, more study is needed to rigorously assess the impact of biometrics on patient health outcomes, preferably via a controlled trial.

Acknowledgments. We are indebted to Sandeep Ahuja and Michael Paik for their contributions to the design and implementation of the biometric terminal. We are also very grateful to all the counselors, providers, and staff at Operation ASHA, who made this study possible.

References

1. Dell, N., Vaidyanathan, V., Medhi, I., Cutrell, E., Thies, W.: "Yours is Better!" Participant Response Bias in HCI. In: CHI (2012)
2. Farmer, T.: Factors influencing adherence to tuberculosis directly observed therapy: A review of the literature. Tech. rep., Toronto Public Health (2005)
3. Krueger, K., Ruby, D., Cooley, P., Montoya, B., Exarchos, A., Djojonegoro, B., Field, K.: Videophone utilization as an alternative to directly observed therapy for tuberculosis. The International Journal of Tuberculosis and Lung Disease 14(6) (2010)
4. SonLa Study Group: Using a fingerprint recognition system in a vaccine trial to avoid misclassification. Bulletin of the WHO 85(1) (2007)
5. World Health Organization: Global Tuberculosis Control 2009: Epidemiology, Strategy, Financing (2009)
6. McDonald, H.P., Garg, A.X., Haynes, R.B.: Interventions to enhance patient adherence to medication prescriptions. Journal of the American Medical Association 288(22) (2002)
7. Moulding, T.S., Caymittes, M.: Managing medication compliance of tuberculosis patients in haiti with medication monitors. The International Journal of Tuberculosis and Lung Disease 6(4) (2002)
8. Nair, D.M., George, A., Chacko, K.T.: Tuberculosis in Bombay: New Insights from Poor Urban Patients. Health Policy and Planning 12(1) (1997)
9. Osterberg, L., Blaschke, T.: Adherence to medication. New England Journal of Medicine 353 (2005)
10. Paik, M., Samdaria, N., Gupta, A., Weber, J., Bhatnagar, N., Batra, S., Bhardwaj, M., Thies, W.: A Biometric Attendance Terminal and its Application to Health Programs in India. In: ACM Workshop on Networking and Systems for Developing Regions (2010)
11. Rao, M.: UID and Public Health: Specious Claims. eSS/Newsclick Policy Matters, Rao, UID health (February 2011)
12. Serwaa-Bonsu, A., Herbst, A.J., Reniers, G., Ijaa, W., Clark, B., Kabudula, C., Sankoh, O.: First experiences in the implementation of biometric technology to link data from health and demographic surveillance systems with health facility data. Global Health Action 3 (2010)
13. Sterns, A.A., Mayhorn, C.B.: Persuasive Pillboxes: Improving Medication Adherence with Personal Digital Assistants. In: IJsselsteijn, W.A., de Kort, Y.A.W., Midden, C., Eggen, B., van den Hoven, E. (eds.) PERSUASIVE 2006. LNCS, vol. 3962, pp. 195–198. Springer, Heidelberg (2006)
14. UID and Public Health. UIDAI Working Paper, http://bit.ly/qwXXgw
15. Weibel, D., Schelling, E., Bonfoh, B., Utzinger, J., Hattendorf, J., Abdoulaye, M., Madjiade, T., Zinsstag, J.: Demographic and health surveillance of mobile pastoralists in Chad: Integration of biometric fingerprint identification into a geographical information system. Geospatial Health 3(1) (2008)
16. Yu, K.L., Chen, C.C., Chang, W.S., Juma, H., Chang, C.S.: Fingerprint identification of AIDS patients on ART. Lancet 365 (2005)

The Neural Persuasion Model: Aligning Neural Readiness, Perceived Need, and Intervention Strategies

Susan Ferebee[1] and James Davis[2]

[1] Kaplan University
[2] University of Phoenix

Abstract. An increasing body of information is being assembled to understand how persuasive technology can be applied to develop successful persuasive health systems. Both technology and human issues contribute to reduced success of some persuasive systems. Neuroscience research has opened the door to improved understanding of how humans process information during attitude formation, attitude change, and during persuasion attempts. This article presents the Neural Persuasion Model, which delves more deeply into the human component of persuasion. The model draws on current neuroscience research and theories of neural readiness and neural organization to suggest ways in which understanding the neural activity of the brain might close the gap between persuasive technology design and behavioral outcomes, particularly for addiction recovery and other circumstances where neural disorder exists.

Keywords: Persuasion, Attitude, Neuroscience, Persuasive Technology.

1 Introduction

Persuasive technology designs are often unsuccessful [1]. Zhang [2] suggested that the failure of health persuasive technology might be caused by inadequate understanding and analysis of human issues rather than by inadequate technology. Allport [3] wrote of the importance of people neurally organizing incoming stimuli to be neurally ready for attitude formation to occur. Current neuroscience research [4], [5], [6], [7], [8], [9] supports Allport's foresight. Functional magnetic resonance imaging (fMRI) allows neuroscientists to examine the human brain response during persuasion and attitude change. The Neural Persuasion Model (NPM) proposed in this study draws on the Theory of Neuroscience [5] and on Allport's theories of neural readiness and neural organization to describe human factors that influence how a persuasive message is received and processed. The NPM suggests intervention strategies based on varying levels of neural readiness and perceived need to change. Understanding the human brain response during persuasion attempts and attitude change can contribute to greater success in persuasive technology design, particularly when intervention targets the most critical situations of addiction recovery.

M. Bang and E.L. Ragnemalm (Eds.): PERSUASIVE 2012, LNCS 7284, pp. 181–192, 2012.

2 Theoretical Framework

The Theory of Neuroscience [5] provides the framework for the Neural Persuasion Model. Neuroscience ties measurable observations about cognitive behavior to physical brain processes that support these behaviors. Oktar [5] suggests four key neuroscience concepts: 1) consciousness, 2) perception, 3) cognition, and 4) behavior. Consciousness is not observable, and it is not something that a person does or can be educated to perform. Oktar describes consciousness as a natural phenomenon. Each individual has access to their own consciousness, but cannot see the consciousness of others. In a consciousness condition, learning and memory become factors of the neuronal state, influencing perception and behavior change. In this neuronal system, entropy, where neural order declines, suggests a disordered state with reduced predictability [5]. Oktar describes cognition (related to working memory) as "sitting between consciousness and behavior" (para. 5). Neural circuits change and adapt over time, during cognitive functions; therefore, the potential to understand how cognition transforms to behavior is enhanced through neuroscience.

3 Literature Review

3.1 Neuroscience

Direct observation of neural activity in real time and in-vivo is possible through functional magnetic resonance imaging (fMRI) [41]. An fMRI is like magnetic resonance imaging (MRI), a noninvasive procedure using radio waves and a strong magnetic field to create an image of tissue and organs within the body. An fMRI, however, measures how blood flows through the brain rather than looking at tissues and organs. Neurons form neural circuits, and the circuits process certain kinds of information (e.g., motor, logical, or sensory), and as processing occurs in a specific brain region, blood flow to that region increases and can be measured. Experiences of different types will activate specific brain regions and modify neural circuits in different regions [7], and additionally, those circuits will relate other external events occurring simultaneously with the experience, and the relationship between the events is permanent within that neural structure [4]. Since 2002, a base of empirical knowledge has developed about how neural circuits, develop and activate during attitude formation and persuasion. Falk and Lieberman [6] suggest the need for an incremental approach where neuroimaging and behavioral research support, inform, and facilitate one another.

3.2 Attitude and Behavior

Miller [10] described attitude as a mediator sitting between a persuasive stimulus and a behavior change. He suggested that the problem is that there is not a direct, empirical measure of attitude, causing attitude to always be inferred. Additionally, according to Miller, there is not always an empirically observed and measurable correlation between inferred attitude and behavior change. Wittenbrink and Schwarz [11] state that "self-reported responses may reflect the products of higher-order thought

processes (e.g., the attitude judgment) or lay theories of cause and effect, but bear little resemblance to the actual computations that produce them" (p. 126). A persuasive method can be seen and measured as can an existence or absence of a behavior change. What cannot be observed is an attitude, and it could be suggested that attitude forms as an element of consciousness, which according to Oktar [5], is also not visible and not measurable. Further supporting this is the work by Lieberman [19] showing that dissonance-based attitude changes were the result of implicit processes, and that attitude change occurred in amnesia patients who had no memory of inconsistent previous behaviors and attitudes. Self-reported reaction to persuasive communication can serve as a predictor for some behavior change, but it has proven less accurate than scientists expected [20]. The possibility now emerges that there can be conscious processes that are not reported by individuals after they receive a persuasive message, as well as implicit processes that occur but do not register in consciousness, so are not reported. Such unregistered and thus unreported processes may explain the variance among behavior changes not explained in self-reporting mechanisms [21], [22], [23].

3.3 Persuasion

Persuasion is communication that has the intent to reinforce, shape, and alter the response of other people [10]. Miller emphasizes that persuasion is intentional. Response shaping, according to Miller, occurs when individuals are faced with unfamiliar situations or unfamiliar people or objects. Response shaping sometimes occurs through social learning [24], [25] as people often mimic other people's behavior when in new situations and they observe which of the observed model's behaviors produce positive outcomes. Response reinforcement occurs through activities, repeated over time, that support the desired behavior. Response alteration, Miller's final phase of persuasion, occurs gradually and often as a result of prior response shaping and response reinforcement. Response shaping will create a response where one did not exist before, but to alter a response means to create a movement from one response to a different response [26].

Persuasion can be direct (message content) or indirect (peripheral elements) and includes the factors of source, receiver, content, context, medium, and expression style [29], [27], [28]. Hasle and Christensen [29] support Scott's [30] epistemic rhetoric view of persuasion as being in and of itself the process of understanding. In this view, rhetoric is limited in its ability to persuade because human comprehension of any topic is inconsistent [31]. According to Fogg [32], persuasion leads to a change in behavior or attitude. Fogg [1] presents three elements that must occur simultaneously to result in behavior change: motivation, trigger, and ability. Motivation can be extrinsic, intrinsic, or can emerge from dichotomous factors such as appealing to fear of pain or a desire for pleasure. Fogg describes a trigger as something visible and relating to the behavior that is to be changed. The trigger should occur when an individual is both capable and motivated to alter the behavior (Kairos moment). Ability refers to a person having the physical capability and resources, as well as cognitive capability to change [1].

Falk et al. (21) examine the neural basis of persuasion and find that the set of brain regions activated during persuasion are similar to those activated when people try to

understand what another is thinking and are also the same as those activated during social cognition tasks. Early neuroscience findings suggest that overlapping neural regions are activated for attitude alteration, persuasion, change in intent, and behavior change. These authors suggest that behavioral and neuroimaging researchers work iteratively, continuing to add to the larger body of knowledge in persuasion and attitude change.

3.4 Persuasive Technology

The field of persuasive technology was preceded by the work of Reeves and Nass [33], who found that individuals responded to and interacted with media and technology just as they would in a social interaction with another person. Fogg [34] extended Reeves and Nass' Media Equation theory and examined what role technology might play in persuasion. Fogg describes computing systems that are designed with the intent to change individual behaviors and attitudes. Fogg's "functional triad" (Chap. 2, p. 1) proposes three ways that computers interact with users: as tools, as media or simulation, and finally as social actors. As social actors, Fogg argues that human persuasion principles of reward and positive feedback, modeling of desired behaviors, and providing socially based support can achieve behavior or attitude alteration. Oinas-Kukkonen and Harjumma [35] extend Fogg's work on persuasive computers and discuss persuasive systems, which combine information systems and computer software. In developing their Persuasive System Design (PSD) model, they identify the contexts needed in persuasive systems as "intent, event, and strategy" (p. 487). Additionally, they describe challenges in persuasive systems in the categories of user issues, persuasion issues, and system issues.

3.5 Persuasive Technology Design

Progress has been made in developing design guidelines for persuasive technology by moving from general concepts to the more specific PSD model [35] that is carefully built on the body of knowledge surrounding persuasive technology including Fogg's [34] triad of computer as tool, media, and social actor. Within the PSD framework are Fogg's seven tools, three media environments, and five social cues. The PSD Model provides a framework that supports dialogue, main task, social factors, and credibility in order to facilitate persuasion [35]. Lykke [36], in applying the PSD model to persuasive technology evaluation, suggested the continuing need to identify the persuasion intervention locus. In spite of these developments in persuasive technology design, the "unpredictable comprehension" [29] of a message by the user remains a problem.

The entire persuasion process is relevant [29], and a simple understanding of self-reported or observed response to individual stimuli and persuasive techniques is not enough [37], [21], [38], [10], [11]. Hovland [39] felt it was so important to understand how the brain responds to stimuli that he spent years programming computers to simulate the human brain responses, explaining that techniques that we think we understand often require further clarity that can then lead to deeper understanding and precision. The introduction of fMRI in the early 1900s at last gave us a tool to pursue

many of the concerns of early theorists like Hovland. The technology of fMRI allows for the application of a neural lens to persuasion. Such an application can lead to an improved precision and understanding of persuasion. Neuroscience may provide answers through a deeper understanding of how learning, decision-making, and self-awareness develop [40].

4 The Neural Persuasion Model

Neural readiness is the basis of all persuasion [3]. Without it, persuasion cannot occur; with it, and combined with a persuasive stimulus, persuasion is possible. The word 'ready' derives from the Old English *rædde* which goes back further to the Germanic *raithjō* meaning arranged, therefore prepared [43]. Allport used the term neural readiness when he defined attitude as a "neural state of readiness, organized through experience, exerting a directive or dynamic influence upon an individual's response to all objects and situations with which it is related" (p. 810). Neural readiness suggests a brain that is prepared to accept and process a persuasive stimulus, with neurons arranged through experience, and without negative external influence. In persuasion attempts, a person's immediate state of neural readiness (which we call neural state in existence (NSIE) reflects a complex matrix of potential neural engagements that will influence how a persuasive stimulus is received. If we want to persuade people to stop a behavior, they must be prepared to accept that stimulus. In other words, a person who has never been educated about the risks of a behavior will be less neurally ready to accept a persuasive message regarding stopping the behavior.

There are factors that can disorder a person's NSIE [5] and therefore affect their ability to receive a persuasive stimulus. These include, but are not limited to: 1) physiological factors like chronic pain [44], substance abuse [45], sleep deprivation [46]; 2) economic factors such as poverty [46] ; 3) external factors like threats or emergency situations [48]; and 4) emotional factors [51] such as grieving. When any of these factors are in a negative state, a person's neural readiness is impaired. For example, fMRI brain scans demonstrate that people with chronic pain have abnormally high and constant activity in an area of the brain related to emotion. This differs from a normal brain where incoming stimuli generates brain activity in many areas throughout the brain [44]. Another example comes from a study, also using fMRI that shows that sleep deprivation (SD) resulted in changed decision-making patterns moving from a pattern of loss-minimization in non-SD participants to a gain-seeking pattern in SD participants [46].

We propose that a neural state in existence is a continuum, but for purposes of research and evaluation, four neural states in existence can be defined. These are:

1. positive NSIE - a neural state when the brain is prepared for and is voluntarily focused on the desired stimulus, and the focus is not hindered by negative physiological, economic, emotional or external factors
2. neutral NSIE - a neural state when there is no particular focus on the desired stimulus, but there are no severe distractions or negative physiological, economic, emotional, or external factors

3. negative NSIE
   a. a neural state when the person is neurally prepared by experience or education, but this state is mediated by one or more of the negative physiological, economic, emotional, or external factors (e.g., person has chronic pain, is sleep deprived, addicted, or in an emergency situation).
   b. a neural state when one or more of the negative physiological, economic, emotional, or external factors are present; and additionally, the person is not prepared through education or experience.

In addition to an individual's NSIE, a factor that interacts with or is mediated by neural readiness is the strength of the individual's perceived need (PN). Need represents a requirement for something because it is perceived as essential or very important. The word need is derived from *nauthiz*, whose non-Germanic origins are necessity and distress, and from German, *not,* meaning danger and emergency [43]. In developing the NPM, we rely on the definition of need that emerges from this sense of urgent necessity. Bandura [24] discusses the relevance of this definition to persuasion when he suggests that a prerequisite to behavior change is increased awareness of a serious threat such as illness. A problem must be identified so that an individual perceives a need for behavior change [50]. This perceived need will interact with or be mediated by the NSIE. A person's persuadability is influenced by their perceived need and NSIE at the moment of receiving a persuasive stimulus.

The NPM suggests five approaches that can be applied to persuasive technology design based on the various possible combinations of NSIE and PN: 1) pre-persuasion-- provide experience that builds neuronal paths that will influence future persuasion response. Identify a problem that generates or exposes the perceived need [4] [5] ; 2) perception shaping [10]-- provide a persuasive message aimed at changing the perception that a need exists; 3) reinforce [34], [10]-- provide persuasive messages that reinforce existing experience and knowledge; 4) reduction [34] --make a complicated or difficult task easier (reducing the number of signals the brain must process); and kairos [34]-- provide timely persuasive messages that arrive within a shifting context of disordering factors to take advantage of moments of kairos when disorder is minimized. For instances, smokers are always disordered to a degree, but can be less disordered when their nicotinic receptors are saturated [49].

Because this model addresses persuading people in states of neural disorder, it suggests a need for persuasive technology that can take advantage of moments of greater neural order to deliver persuasive messages. Kairos is therefore, defined for this study as "a passing instant when an opening appears which must be driven through with force if success is to be achieved" [50].

Tables 1 through 3 describe the NPM and provides a matrix of perceived need for change, neuronal response, education/experience, and possible approaches that can guide persuasive technology designers through a consideration of the various combinations. These are suggestions to stimulate further thinking, and clearly, empirical research will be required to address the optimal persuasion approach for each category.

Table 1. High PN across varying NSIE with proposed persuasive approach

	Not Neurally Disordered		**Neurally Disordered**	
	Positive NSIE	Neutral NSIE	Negative NSIE –a	Negative NSIE-b
High PN Urgent and Important	Individual is educated or experienced on topic and has no neurally disordering factors	Individual is not educated on the topic and has no neurally disordering factors	Individual is educated or experienced on topic but does have neurally disordering factors.	Individual is not educated or experienced on the topic, and has neurally disordering factors.
Persuasive Approach	**All techniques**	**Shaping**	**Kairos Reduction**	**Shaping Kairos Reduction**

Individuals with a high perceived need, fully educated on the topic in question, and with no neurally disordering factors are likely to respond to a wide variety of persuasive techniques. For those individuals who are not neurally disordered but who lack education/experience in a topic, shaping of the need is a reasonable strategy because until the need is understood, the person is unlikely to respond to persuasive stimuli (51). For all levels of NSIE that represent the presence of a neural disordering factor (e.g., addiction), reduction is suggested to reduce the cognitive load required by the persuasive stimulus. Although a smoker is educated on the risks of smoking, creating a high perceived need and motivation to stop, the neural disordering caused by nicotine addiction can alter how a persuasive message is perceived as well as the person's ability to process the message or take action.

Table 2. Medium PN across varying NSIE with proposed persuasive approach

	Not Neurally Disordered		**Neurally Disordered**	
	Positive NSIE	**Neutral NSIE**	**Negative NSIE-a**	**Negative NSIE-b**
Medium PN Important Not Urgent	Individual is educated or experienced on topic and has no neurally disordering factors	Individual is not educated on the topic and has no neurally disordering factors	Individual is educated or experienced on topic but does have neurally disordering factors.	Individual is not educated or experienced on the topic, and has neurally disordering factors.
Persuasive Approach	**Reinforcing**	**Shaping**	**Reinforcing Reduction Kairos**	**Pre-persuasion Shaping Kairos Reduction**

When a person feels that a need for change is important but not urgent, motivation is reduced. A person in this category who is educated about the topic and has no neurally disordering factors will require that the understood need for change be reinforced. For a person who is not educated on the topic (again, perhaps the risks of smoking), some shaping messages focused on education may be necessary to create

the required need, and then once the perceived need is elevated to high, other techniques can be employed. A person with a medium perceived need that is also neurally disordered will require the same reinforcement and possible shaping techniques. However, reduction and Kairos are suggested as methods to compensate for the reduced cognitive ability caused by the disordering factor.

Table 3. Low PN across varying NSIE with proposed persuasive approach

	Not Neurally Disordered		**Neurally Disordered**	
	Positive NSIE	**Neutral NSIE**	**Negative NSIE-a**	**Negative NSIE-b**
Low PN Not Important and Not Urgent	Individual is educated or experienced on topic and has no neurally disordering factors	Individual is not educated on the topic and has no neurally disordering factors	Individual is educated or experienced on topic but does have neurally disordering factors.	Individual is not educated or experienced on the topic, and has neurally disordering factors.
Persuasive Approach	**Shaping**	**Shaping**	**Shaping**	**Pre-persuasion Monitor Surveillance**

A person who has a low perceived need for behavior change and who is classified as NSIE-b (neurally disordered and with no education about the topic) is not ready to accept a persuasive stimulus. An example is a person at a bar who is drunk and wants to drive home. Driving home is not perceived as a threat due to the neurally disordering effect of intoxication, and getting home is perceived as a reward. Persuasion cannot occur in this state. The use of built-in automobile breathalyzers provides surveillance and monitoring that prevents a person from driving when intoxicated, and will contribute to pre-persuasion education about the consequences of drinking. On the other hand, if this same person is at the bar but does not drink (and therefore is not neurally disordered), yet still feels a low need to stop drinking, the breathalyzer on the car will allow the ignition to start, and the reward will begin to shape attitude and provide education.

Many persuasive technologies exist to help people stop smoking or stop other forms of addiction. Räisänen, Oinas-Kukkonen and Pahnila [54] asked the question of when the opportune moment to persuade a person to quit smoking might be. The NPM provides insight here. A study by Kolb and Wisha [51] supports the NPM in that the researchers found that smokers who have a lower perceived need to quit, respond significantly less favorably to all types of cessation ads. Additionally, cessation ads were less effective for very heavy smokers, and it can be posited that the NSIE is more disordered with higher levels of nicotine in the system. The NPM suggests that optimal timing of the persuasive message might be counter-intuitive. If you have a person highly motivated to stop smoking, the receipt of a message during a craving moment is not likely to be processed as intended because of disorder to NSIE as brain regions associated with compulsive behavior, arousal, and episodic memory are active during that craving [49]. On the other hand, immediately after smoking, individuals experience a sense of well-being and higher energy through the release of

dopamine [54]. While this sense of well-being is nicotine induced and also represents a disordering of NSIE, it is a less negative disordering and suggests that delivering a persuasive message during smoking and within 15 minutes after smoking will have a greater chance of that message being processed.

4.1 Use of NPM to Classify Smokers and Suggest Persuasive Techniques

The goal of persuasion and persuasive technology is twofold: identify the needs of users and identify the optimal persuasive message and technology that meet those needs, thus leading to the desired behavior change. Mediating both of these will be the neural readiness of the participant. A smoker's perceived need for behavior change can be high, medium, or low as described in the NPM. Whether the person will successfully stop smoking will be influenced by that person's NSIE. Disordering factors to neural readiness exist in many people at many times and are often unknown by those trying to persuade. However, with smokers, the act of being a smoker creates a neurally disordered state. Additionally, as a smoker abstains, it is shown that their brain response to smoking cues increases significantly, creating a further disordering with regard to attention and decision-making as the cessation goal is achieved [52]. This has to be considered in a smoking cessation program. Neural readiness has two elements – the state of the brain activity and the state of being educated or informed about a topic. The education and information plays a major role in the creation of the need state. For example, a smoker with a previously low perceived need to stop smoking who receives information from a physician about x-ray evidence that shows pulmonary dysfunction, now has a high perceived need to stop smoking, but is not well educated on the topic and is neurally disordered due to nicotine addiction. Using the NPM, this might be recognized as the Kairos moment (in the doctor's office) to begin education and provide reduction techniques as part of a no smoking program. Shaping has importance in developing an attitude change about the danger of smoking, and reduction is necessary to simplify the required cognitive processing while the person is neurally disordered by nicotine addiction. Considering the matrix as a whole, the NPM matrix for smokers shrinks to include only the right side quadrants because of the smokers' pre-existing neurally disordered state. This sharpens the persuasion focus.

5 Discussion

This study presents a neural persuasion model to better predict the outcomes of persuasive attempts using persuasive technology. The focus on the individual being targeted starts at the foundation of persuasion: the neural state of readiness [3]. The cells of the NPM matrix suggest individual differences that might account for the variable success rate of persuasive attempts. The specifications of each cell may well provide the persuasive designer with information that closes the gap between intent and behavioral outcomes. If one accepts the logic that the state of neural readiness will determine success or failure, then the exploration of each cell's predictive ability

will be valuable. Applying the neural lens to persuasion expands the understanding of factors that mediate and limit motivation and ability. A person could be motivated to exercise more and have the physical ability, but a mental state of depression would mean that neural readiness is not present. Clearer specifications of what needs to go into a persuasion strategy are likely to be useful. Greater precision and accuracy with respect to targeted behaviors seems a reasonable outcome. The application of the NPM to a variety of other outcomes for persuasion technology is straightforward. The linkage of persuasion to neural readiness is a connection that ought to have a favorable outcome.

6 Conclusion

The need to understand the neural readiness of a person when an attitude is being constructed or changed has existed since the earliest persuasion studies. With the new imaging technology available, we can now address that need. The NPM presented here suggests new ways to connect persuasive technology with behavior change. The goal is to test all twelve cells of the NPM matrix and to test them across as many desired persuasive outcomes as possible.

References

1. Fogg, B.J.: Creating Persuasive Technologies: An Eight-step Design Process. In: Persuasive 2009, Claremont, CA, USA, April 26-29 (2009)
2. Zhang, J.: From Topological Perception to Distributed Cognition. Visual Cognition 12, 662–664 (2005)
3. Allport, G.: Attitudes. In: Murchison, C. (ed.) A Handbook of Social Psychology, pp. 789–844. Clark University Press, Worcester (1935)
4. Aimone, J., Wiles, J., Gage, F.: Computational Influence of Adult Neurogenesis on Memory Encoding. Neuron 6(2), 187–202 (2009)
5. Oktar, N.: Theory of Neuroscience. Journal of Neurological Sciences 23(3), 155–158 (2006)
6. Falk, E., Lieberman, M.: The neural bases of attitudes, evaluations, and behavior change. In: Kruger, F., Grafman, J. (eds.) The Neural Basis of Human Belief Systems, Psychology Press (2012)
7. Ramachandran, V.: The Tell-Tale Brain. W. Norton & Company, New York (2011)
8. Small, G., Vorgan, G.: IBrain. Harper Collins ebooks (2009)
9. Wang, L., Conner, J., Rickert, J., Tuszynski, M.: Structural Plasticity with Highly Specific Neuron Population Identifies a Unique Parcellation of Motor Learning in the Adult Brain. Proceedings of the National Academy of Sciences (in press)
10. Miller, G.: On Being Persuaded: Some Basic Distinctions. In: Roloff, M., Miler, G. (eds.) Persuasion: New Directions in Theory and Research. Sage Publications (1980)
11. Wittenbank, L., Schwartz, N.: Implicit Measures of Attitude. The Guilford Press, New York (2007)
12. Festingr, L., Carlsmith, J.: Cognitive Consequences of Forced Compliance. Journal of Abnormal and Social Psychology 58(2), 203–210 (1959)

13. Festinger, L.: A Theory of Cognitive Dissonance. Stanford University Press, Stanford (1957)
14. Hovland, C., Janis, I., Kelly, H.: Communications and Persuasion: Psychological Studies in Opinion Change. Yale University Press, New Haven (1953)
15. Eagly, A., Chaiken, S.: The Psychology of Attitudes. Harcourt, Brace, Ft. Worth, TX (1993)
16. Katz, D.: The Functional Approach to the Start of Attitudes. Public Opinion Quarterly 24, 163–204 (1960)
17. Ajzen, I., Fishbein, M.: Understanding Attitudes and Predicting Social Behavior. Prentice-Hall, Englewood Cliffs (1980)
18. Ajzen, I.: The Theory of Planned Behavior. Org. Behav. Hum. Decis. Process. 50, 179–211 (1991)
19. Lieberman, M., Ochsner, K., Gilbert, D., Schacter, D.: Do Amnesiacs Exhibit Cognitive Dissonance Reduction? The Role of Explicit Memory and Attention in Attitude Change. Psychological Science 12, 135–140 (2001)
20. Webb, T., Sheeran, P.: Does Changing Behavioral Intentions Engender Behavioral Change? A Meta-Analysis of the Experimental Evidence. PsycholBull. 132, 249–286
21. Falk, E., Berkman, E., Mann, T., Harrison, B., Lieberman, M.: Predicting Persuasion-induced Behavior Change from the Brain. Journal of Neuroscience 30, 8421–8424 (2010)
22. Lieberman, M.: Social Cognition Neuroscience: A Review of Core Processes. Annual Review of Psychology 58 (18.11) (2007)
23. Ochsner, K., Lieberman, M.: The Emergence of Social Cognition Neuroscience. American Psychologist 58(8.11) (2001)
24. Bandura, A.: Perceived Self-efficacy in the Exercise of Personal Agency. Journal of Applied Sport Psychology 2, 128–163 (1990)
25. Burgoon, J., Burgoon, M., Miller, G., Sunnafrank, M.: Learning Theory Approaches to Persuasion. Human Communications Research 7, 161–179 (1981)
26. Stiff, J., Mongeau, P.: Persuasive Communication. Guilford Press, New York (2003)
27. Petty, R., Cacioppa, J.: Communication and Persuasion: Central and Peripheral Routes to Attitude Change. Springer, New York (1986)
28. Powers, P.: Persuasion and Coercion: A Critical Review of Philosophical and Empirical Approaches. HEC Forum 19(2), 125–143 (2007)
29. Hasle, P., Christensen, A.: Persuasive Design. In: Kelsey, S., Amant (eds.) Handbook of Research on Computer Mediated Communication, pp. 283–296. Information Science Reference, Hersey (2008)
30. Scott, R.: On Viewing Rhetoric as Epistemic. Central States Speech Journal 18, 9–16 (1967)
31. Nielsen, K.: An Idea Critic: Ciceronian Rhetoric and Contemporary Criticism. Peter, Bern (1995)
32. Fogg, B.J.: Mass Interpersonal Persuasion: An Early View of a New Phenomenon. In: Oinas-Kukkonen, H., Hasle, P., Harjumaa, M., Segerståhl, K., Øhrstrøm, P. (eds.) PERSUASIVE 2008. LNCS, vol. 5033, pp. 23–34. Springer, Heidelberg (2008)
33. Reeves, B., Ness, C.: The Media Equation: How People Treat Computers, Television, and New Media Like Real People and Places. Cambridge University Press, London (1996)
34. Fogg, B.J.: Persuasive Technology: Using Computers to Change What We Think and Do. Morgan Kaufmann Publishers, New York (2003)
35. Oinas-Kukkonen, H., Harjumaa, M.: Persuasive Systems Design: Key Issues, Process Model, and System Features. Communications of the Association for Information Systems 24, 485–500 (2009)

36. Lykke, M.: Persuasive Design Strategies: Means to Improve the Use of Information Organization and Search Features in Web Site Information Architecture. In: Proceedings ASIST Special Interest Group pm Classification Reserach 20th Workshop, Vancover, B.C. (2009)
37. Berns, G., Moore, S.: A Neural Predictor of Cultural Popularity. Journal of Consumer Psychology (in press)
38. Krosnick, J., Judd, C., Wittbank, B.: The Measurement of Attitudes. In: Albarraceni, D., Johnson, B., Zama, M. (eds.) The Handbook of Attitudes, pp. 21–76. Erlbuam, Maheahah (2005)
39. Hovland, C.: Complete Simulation of Thinking. American Psychologist 15(11), 687–693 (1960)
40. Churchland, P.: The Impact of Neuroscience on Philosophy. Neuron 60(3), 409–411 (2008)
41. Lichtman, J., Smith, S.: Seeing Circuits Assemble. Neuron 60(3), 441–448 (2008)
42. Marois, R., Ivanoff, J.: Capacity Limits of Information Processing in the Brain. Trends in Cognitive Processing 9(6), 296–305 (2005)
43. OED, http://www.oed.org (retrieved from November 11, 2011)
44. Pujol, J., López-Solà, M., Ortiz, H., Vilanova, J., Harrison, B., Yücel, M., Soriano-Mas, C., Cardoner, N., Deus, J.: Mapping Brain Response to Pain in Fibromyalgia Patients Using Temporal Analysis of fMRI. PLoS ONE 4(4), e5224 (2009)
45. Risinger, R., Salmeron, B., Ross, T., Amen, S., Sanfilipo, M., Hoffmann, R., Bloom, A., Garavan, H., Stein, E.: Neural Correlates of High and Craving During Cocaine Self-administration Using BOLD fMRI. Neuroimage 26, 1097–1108 (2005)
46. Venkatraman, V., Huettel, S., Chuah, L., Payne, J., Chee, M.: Sleep deprivation biases the neural mechanisms underlying economic preferences. J. Neurosci. 31(10), 3712–3718 (2011)
47. Silverman, M., Muennnig, P., Liu, X., Rosen, Z., Goldstein, M.: The Impact of Socioeconomic Status on the Neural Substrates Associated with Pleasure. Open Neuroimag J. 3, 58–63 (2009)
48. Pichon, S., Gelder, B.D., Grezes, J.: Threat Prompts Defensive Brain Responses Independently of Attentional Control. Cerebral Cortex (2011) (in press)
49. Brody, Mandelkern, London, Childress, Lee, Bota, Ho, Saxerna, Baxter, Madsen, Jarvik: Brain Nicotinic Acetylcholine Receptor Occupancy: Effect of Smoking a Denictotinized Cigarette. International Journal of Neuropsychopharmacology (published online)
50. White, E.: Kaironomia. Cornell University Press (2008)
51. Kolb, B., Wisha, I.: Fundamentals of Neuropsychology. Worth Publishers (2008)
52. McClernon, F., Kozink, R., Lutz, A., Rose, J.: 24-hour smoking abstinence potentiates fMRI BOLD activation to smoking cues in cerebral cortex and dorsal striatum. Psychopharmacology 204, 25–35 (2009)

Analyzing the Persuasion Context of the Persuasive Systems Design Model with the 3D-RAB Model

Isaac Wiafe, Muna M. Alhammad, Keiichi Nakata, and Stephen R. Gulliver

Informatics Research Centre
Henley Business School
Whiteknights
University of Reading
Reading, UK
{i.wiafe,m.m.s.alhammad}@pgr.reading.ac.uk,
{k.nakata,s.r.gulliver}@henley.reading.ac.uk

Abstract. Research into design methodology is one of the most challenging issues in the field of persuasive technology. However, the introduction of the Persuasive Systems Design model, and the consideration of the 3-Dimensional Relationship between Attitude and Behavior, offer to make persuasive technologies more practically viable. In this paper we demonstrate how the 3-Dimensional Relationship between Attitude and Behavior guides the analysis of the persuasion context in the Persuasive System Design model. As a result, we propose a modification of the persuasion context and assert that the technology should be analyzed as part of strategy instead of event.

Keywords: Persuasive Technology, Behavior Change, Captology, Systems Design.

1 Introduction

The Persuasive System Design (PSD) model [1] is currently the only comprehensive and systematic framework for developing and evaluating persuasive systems. It helps to structure thinking and provides maps for selecting system features and requirements, and thus has been successfully applied in developing and evaluating a number of technologies [2, 3]. However, it does not provide tools that can be used during the analysis of the persuasion context, which is crucial to the successful use of the model.

To overcome this limitation, we applied the 3-Dimensional Relationship between Attitude and Behavior (i.e. the 3D-RAB model [4]) to the PSD model. We begin by presenting an overview of the PSD model and this is followed by a description of the 3D-RAB model. We show how the 3D-RAB model provides a useful guide in analyzing *the intent*, *the event*, and *the strategy* of design, and argue that the *technology context* of the PSD model should be studied as a part of design strategy, instead of it being part of *event* analysis.

M. Bang and E.L. Ragnemalm (Eds.): PERSUASIVE 2012, LNCS 7284, pp. 193–202, 2012.

2 The PSD Model

The Persuasive Systems Design (PSD) model [1, 5] assists the development and evaluation of persuasive technology and consists of three main phases. It is based on the principle that before a successful persuasive system is designed, the designer needs to i) understand the key issues behind the development, ii) analyze the persuasion context, and iii) ensure that the systems demonstrate a specific set of qualities or features.

It argues that understanding the main issues behind persuasive systems is a vital step in designing persuasive technology [1], which is followed by an analysis of the persuasion context. Here, there is the need to study *the intent, the event*, and *the strategy* [6]. It is also explained that the persuasion context is more useful when studied through the idea of cognitive inconsistency [5]. The persuasion context of the PSD model is presented diagrammatically in Figure 1.

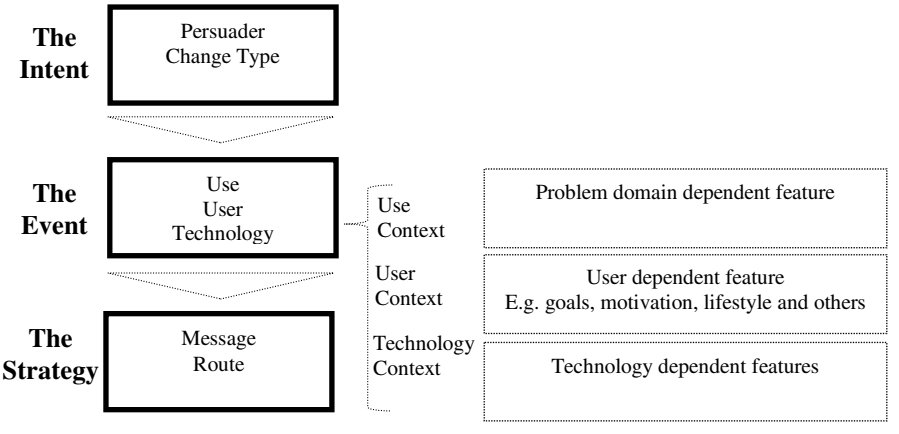


Fig. 1. The persuasion context of the PSD model [1, p. 489]

The analysis of the persuasion context is an essential step needed to promote attitude or behavior change. It starts with studying and identifying *the intent* that comprises the persuader and the change type. There are three main types of persuaders: those who create the interactive technology (endogenous); those who distribute the technology to others (exogenous); and the very person who uses the interactive technology (autogenous) [7]. It also states that there are two main types of change namely behavior change and attitude change.

The next step of analyzing persuasion context is the understanding of *the event*, which aims to analyze the *use context*, *user context* and the *technology context*. The designer needs to study and understand the domain specific issues within a specific *use context*. Oinas-kukkonen and Harjumaa [1] argue that there is a critical need to identify user specific issues in relation to the design of the systems, i.e. consideration should be given to individual differences, in terms of specific user information

processing approach and the target goals [6]. The third issue in event analysis is the consideration of *technology context*, which deals with the strengths and weaknesses of the technology that would be used for designing the system.

The analysis of *the strategy* is another important step in analyzing the persuasion context. At this stage, the designer attempts to identify *the route* that will promote persuasion. More importantly, focus is on the selection of peripheral or central route as proposed by Petty [8, 9]. The routes could be direct or indirect depending on the user characteristics and ability to assess the meaning of persuasive messages.

The analysis of persuasive system features is also a major factor that is considered in the design of persuasive technology. Accordingly, the designer needs to ensure that the system demonstrates a set of non-functional requirements which includes: *dialogue support; social support; systems credibility support* and *primary task support*. In each of the features, there is a set of techniques which can be used [1].

PSD is a high level model, and hence designers may find it difficult to apply it effectively since the information it provides is not sufficient to guide the designer concerning how to perform / apply the activities mentioned in the model. This limits its full potential use. Moreover, as mentioned above, PSD considers the analysis of the technology context as part of the event (see Fig. 1.), which implies that technology is preselected before the strategy is analyzed. However, it seems more appropriate for the technology to be selected as part of the strategy, in order to conform to the message and the route.

In the following section we argue that an analytic model of attitude and behavior the 3D-RAB, can be used to: i) *analyze the change type*; ii) *the user context* iii) *the use context* and iv) *the route*.

3 The 3D-RAB Model

The 3D-RAB model [4] captures the 3-Dimensional Relationship between Attitude and Behavior. It supports the persuader, or the system's designer, to analyze the user or the persuadee so as to adopt or adapt a strategy to support effective persuasion. The 3D-RAB model is used for categorizing users' state of cognitive dissonance based on their current behavior (CB), attitude towards a target behavior (ATTB), and attitude towards changing their current negative behavior or maintaining their positive behavior (ATCMB). The 3D-RAB model advocates that target behavior should be a simple unit of behavior that is measureable in terms of it being positive or negative. A user's current behavior is thus considered 'positive' whenever it is the same as the target behavior and negative otherwise. In total, the 3D-RAB consists of 8 states (see Fig. 2), and each state is characterized by specific attitude and behavior of a user. Also, it continues that the designer must analyze the user's immediate and external environmental factors, i.e. factors that may promote or impede the facilitation of behavior change. Particularly, attention should be given to planned and unplanned attitude and behavior changes [10].

Moreover, the 3D-RAB model extends the cognitive dissonance theory[1][11] and argues that there are variations in levels of dissonance that can be classified into i) strong, ii) moderate, iii) weak or iv) absent of, or No dissonance. See table 1.

Table 1. Cognitive state changes based on the strength of dissonance [4]

State	CB	ATTB	ATCMB	Dissonance Level	Stability of State	Natural State Transition Tendency	Targeted State for Persuasion
1	**+**	**+**	**+**	No	Stable (+)	1	1
2	**+**	**+**	**-**	Weak	Unstable (+)	1	1
3	**+**	**-**	**+**	Moderate	Unstable (-)	7	1
4	**+**	**-**	**-**	Strong	Unstable (-)	8	2 or 3
5	**-**	**+**	**+**	Strong	Unstable (+)	1	1
6	**-**	**+**	**-**	Moderate	Unstable (-)	8	2 or 5
7	**-**	**-**	**+**	Weak	Unstable (-)	8	3 or 5
8	**-**	**-**	**-**	No	Stable (-)	8	4 or 6 or 7

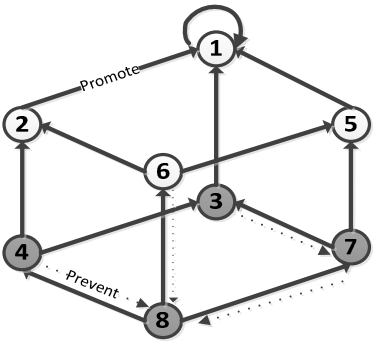


Fig. 2. Diagrammatic representation of possible transitions in 3D-RAB [4]

Wiafe et al. [4, 10], states that there are various paths or routes that can be used to achieve persuasion, and thus it is necessary for the designer to identify the most appropriate and effective route for a user to take in order to achieve effective and efficient behavior change. If a user is found in state 8 (i.e. they have a negative attitude towards behavior, negative behavior and negative attitude towards change), there are six possible routes that can be used to guide persuasion, assuming the aim of the designer is to move a user to state 1. Table 2 contains possible routes from state 8 to 1. In such a scenario, each route needs to be studied carefully, and considerations should be given to the environmental factors that will promote a particular route of persuasion to ensure that the most effective path towards positive change is made.

[1] The cognitive dissonance theory states that an individual experiences an unpleasant feeling when his or her actions are not consistent with his or her beliefs.

Table 2. A list of possible routes from state 8 to 1 in 3D-RAB

	Possible Routes
1.	8→7→3→1
2.	8→7→5→1
3.	8→4→2→1
4.	8→4→3→1
5.	8→6→2→1
6.	8→6→5→1

Another use of the model is its ability to monitor a user's progressive change during a persuasive intervention. This can be achieved via the use of questionnaire, and allows us to identify before and after states of users.

4 Analyzing the Persuasion Context with 3D-RAB

When using the PSD model, the analysis of the persuasion context is considered to be challenging, as it forms the core of the design analyses process. For example, in the development of Fit4life [2], the designers claimed to have applied the PSD model, yet they did not provide sufficient information on the process used to perform the analysis. Räisänen et al. [12] explained that the PSD model is more useful when the persuasion context is studied using a supporting behavior change theory. This supports our application of the 3D-RAB model on PSD. To demonstrate how the persuasion context can be analyzed by applying the 3D-RAB, the following sections expand the process, using the scenario of designers developing a smoking cessation application for in-patients at a hospital.

4.1 The Intent

As mentioned earlier, the analysis of *the intent* comprises two main activities. The identification of the persuader is relatively straightforward since the designer can classify the persuader based on Fogg's classification [13]. Although one can argue that it should also be easy for a designer to state the type of change being addressed, i.e. as being behavior or attitude change, most designers of persuasive technologies do not always assess the type of change being achieved. Sterns and Mayhorn [14], for example, claimed that their systems aimed to change attitude towards behavior, but observation showed that the system actually changed behavior, which does not automatically imply a change in attitude towards behavior. Similarly, even though one may argue that changing attitude towards a behavior can lead to a behavior change, as in the case of the theory of planned behavior [15], designers should be specific on the design intent, to enable a meaningful evaluation of the success of their design.

By applying the 3D-RAB model to the PSD model, we identified that although there are two main types of change to be considered, a further distinction should be

made when focusing on behavior change. Behavior change cannot be considered in isolation without attitude (attitude towards the behavior and attitude towards the change or maintaining the behavior); a strong relationship that is well evidenced in literature [15-18].

In the hospital example, the designer can be considered as the persuader, and thus classified as both exogenous and endogenous. Although the aim of the system is to make in-patients quit smoking, there is a need to have specific objectives that can be achievable and measurable. A total change in behavior may be unfeasible. Accordingly, specifications should be defined, based on user assessment, relating to how many cigarettes the system will encourage the user not to smoke. This approach of behavior definition enables the designer to have a specific target, which can be easily addressed; this can then be linked into the behavior wizard proposed by Fogg and Hreha [19].

According to the 3D-RAB model, consideration should also be given to the user's current attitude towards smoking as well as his/her attitude towards change to, or maintenance of, positive behavior. The benefit of defining these parameters is that it enables the designer to consider scenarios that might escape a designer during *the intent* analysis. For instance, there is a possibility that some in-patients, who smoke during their daily lives, may not wish to smoke when they are at the hospital; whereas ex-smoker in-patients may revert to smoking as a result of being admitted to hospital. The former may do so due to increased cognitive dissonance, and the latter may be using smoking as a personal approach to combat stress, even though they might be experiencing cognitive dissonance as a result of their actions.

4.2 The Event

In analyzing the user context, the 3D-RAB serves as a tool that facilitates user analysis, since it categorizes users into 8 states based on variations in cognitive dissonance; and as such provides useful information about users in a particular category. It is envisaged that this information will enable designers to be more specific in the selection of effective persuasive routes, and message content that will be appropriate for effective persuasion. In addition, the analysis of the *use context* is simplified since it is expected that the distribution in target user's cognitive state is domain specific. For example, most smokers accept the strong medical evidence that smoking is significantly detrimental to health, which means that patients are more likely to be found in state 5, rather than state 8. This is perhaps not true for persuasive systems used to impact issues where people have a range of diverse opinions and beliefs, often based on the cultural and religious backgrounds, e.g. systems aimed at reducing premarital sex. In such cases, the 3D-RAB can guide designers by measuring the current cognitive state of target users, and determine the possible external factors (such as culture, religion, norms, legislature and other possible planned behavior change) that may promote or impede the intended behavior [10].

By using the 3D-RAB model for analyzing the *use* and *user context*, the designer does not only benefit from a simple, yet more specific method, but also acquire information that can be used to measure the success of change. As the designers can

administer the same questionnaire before and after the implementation, they will be able to identify changes that have occurred in users over time, as a result of defined interaction. Since the 3D-RAB aims at identifying disparities in the three factors, it is effective in identifying related changes in cognition that may lead to change in users' behavior in future.

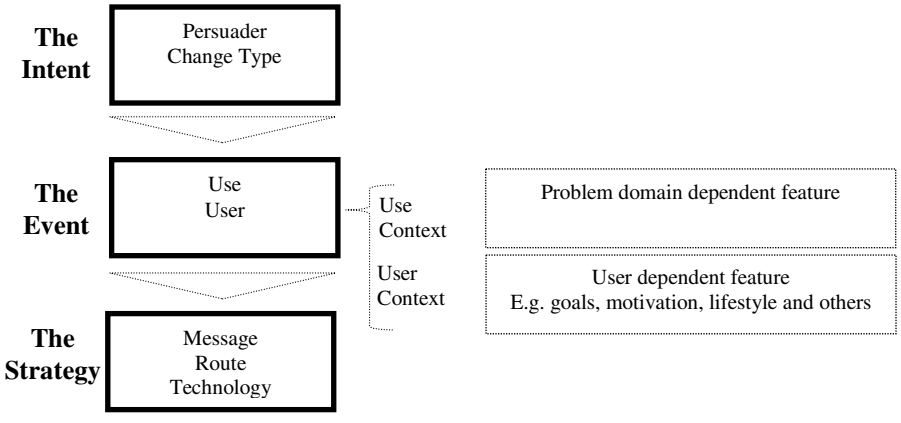


Fig. 3. The modified version of the persuasion context of the PSD model

4.3 The Strategy

The strategy plays a key role in the success of the entire design and we propose three activities instead of the two in the PSD [1, 5]. As shown in Fig. 3, *the strategy* now considers the analysis and selection of technology, rather than being studied in context of *the event* (as in Fig. 1).

By using the 3D-RAB model, the designer is capable of selecting a persuasive *route* for each user group. This selection is based on the information acquired during the analysis of the event, by considering both the user's state, and external influencing factors. Since the 3D-RAB provides information on current states and possible transitions, the designer can choose the most appropriate path to achieve effective persuasion towards the target state. Each path is characterized by specific properties, and thus methods can be selected to implement the intended change.

In some instances, it is appropriate for the designer to focus on using a central route; in others it is appropriate to use peripheral routes as proposed by Petty [8] in the Elaborative Likelihood Model (ELM). The ELM treats persuasion as a cognitive event by targeting persuasive messages to use mental processes of motivation and reasoning to accept or reject messages and it focuses on induced attitudinal changes [8]. It proposes that centrally routed messages targets the creation of long-term changes whereas peripheral messages are normally for short-term changes. Accordingly, centrally routed messages are more effective in attitudinal changes whereas

peripheral messages are more useful in attaining compliance to behavior. Hence by applying the 3D-RAB, a designer will be capable of identifying routes that are appropriate for persuasion. Figure 4 presents a matrix of states transitions and suggested routes for persuasion using the ELM.

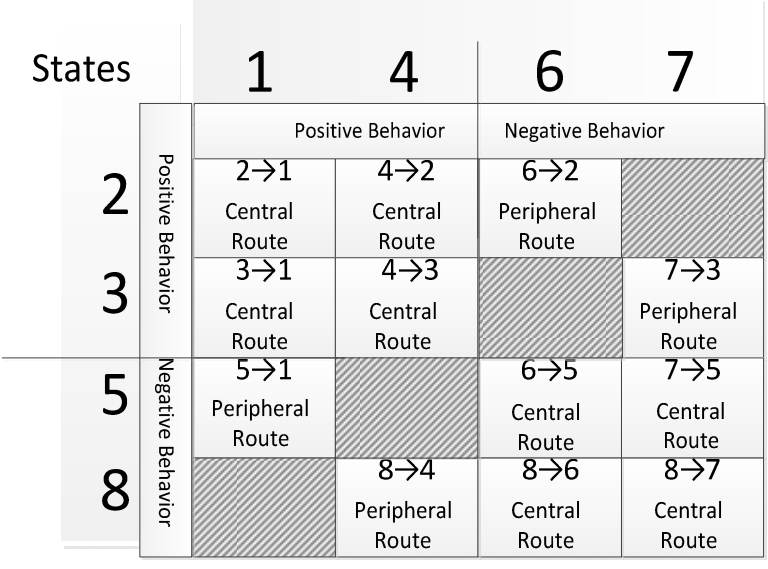


Fig. 4. States transitions and suggested routes for persuasion using the ELM

Consider a user, for instance, in state 5 where both attitude towards behavior and attitude towards change are positive. In such cases, the designer should focus on using the peripheral route to move the user to state 1. This is because the user is already aware of the benefit of him/her changing his/her current behavior hence the designer should focus on cues and triggers that will promote the behavior as compared to the case where both attitudes are negative. Since the 3D-RAB enables the designer to analyze the *use context*, and allows consideration of cultural and other external factors, the designer would be capable of identifying the appropriate message and platform that will promote the target behavior in the selected route.

Based on the analysis above, we deemed it appropriate to consider technology as part of *the strategy*, instead being studied in context of *the event*. Classification of technology as part of the strategy, as in Figure 4, allows the selection of technology for the development of the system, instead of limiting system functionality to ensure the system will work on predefined technology. The designer is accordingly capable of identifying the most appropriate technology for the solution in terms of affordability, ease of use, and acceptance of technology; whilst also considering how this persuasive message and route can be implemented to maximize impact. We acknowledge that there are instances where the designer is limited in the selection of technology, yet these should be defined as strategic limitations.

When considering the scenario of the smoke cessation system for in-patients, if a user is identified to be in state 8, there are 6 possible paths that can be considered (see Table 2). Experiences, for example, may show that in-patients who are not staying at the hospital for a long period of time are more likely to give up smoking if you focus on changing their behavior first. Since the hospital prohibits smoking in the wards, and patients normally smoke outside the building, the designer should consider the various types of change process, messages and technology use that can be employed for persuasion. In this example, routes iii) 8→4→2→1 or iv) 8→4→3→1 may be more suitable. In 8→4→2→1, the designer aims at employing tactics that will stop patients from smoking; and later aim at changing attitude towards behavior, and attitude towards maintaining that change. In 8→4→3→1, attention is given to maintaining the behavior, followed by application of methods to support change in the patients' attitude towards the target behavior. It can be observed that in all the selected paths, both central and peripheral routes are used, however the sequence in which it is presented are different.[2] Each route may be preferable in context of application, yet domain application knowledge cannot be assumed at this time.

5 Conclusion and Future Work

In this paper we have discussed how the 3D-RAB model can be applied to analyze the persuasion context of the PSD model. It was identified that the 3D-RAB is capable of guiding and simplifying the analysis of the persuasion context. Also, we suggested that the technology analysis in the PSD should be considered as part of *the strategy* rather than *the event*. Here, we argued that analyzing technology as a strategy enables the designer to choose technology to suit the message and the route instead of choosing message and route to suite technology.

Finally, it can be observed that the scope of our study was limited to analyzing the persuasion context and even though an explanation was given on how to perform the analysis effectively, no information was provided on specific system features that will support a particular route or path. However, such information will be useful to designers. As such, future studies should focus on effective methods for selecting systems features for specific routes or paths for persuasion. Moreover, we envisage that the potentials of applying 3D-RAB to PSD model extends beyond those highlighted in this paper and thus there is an opportunity for further research to identify other possible applications.

References

1. Oinas-Kukkonen, H., Harjumaa, M.: Persuasive systems design: Key issues, process model, and system features. Communications of the Association for Information Systems (2009)
2. Purpura, S., Schwanda, V., Williams, K., Stubler, W., Sengers, P.: Fit4life: The design of a persuasive technology promoting healthy behavior and ideal weight. ACM (2011)

[2] The example considered the case where there is a need for a complete change and accordingly it was required that both central and peripheral routes were needed to implement a successful change.

3. Segerståhl, K., Kotro, T., Väänänen-Vainio-Mattila, K.: Pitfalls in Persuasion: How Do Users Experience Persuasive Techniques in a Web Service? In: Ploug, T., Hasle, P., Oinas-Kukkonen, H. (eds.) PERSUASIVE 2010. LNCS, vol. 6137, pp. 211–222. Springer, Heidelberg (2010)
4. Wiafe, I., Nakata, K., Moran, S., Gulliver, S.: Considering user attitude and behaviour in persuasive systems design. In: European Conference on Information Systems, Helsinki, Finland (2011)
5. Oinas-Kukkonen, H., Harjumaa, M.: A Systematic Framework for Designing and Evaluating Persuasive Systems. In: Oinas-Kukkonen, H., Hasle, P., Harjumaa, M., Segerståhl, K., Øhrstrøm, P. (eds.) PERSUASIVE 2008. LNCS, vol. 5033, pp. 164–176. Springer, Heidelberg (2008)
6. Oinas-Kukkonen, H.: Behavior change support systems: The next frontier for web science. In: Web Science Conference 2010, Raleigh, NC, USA (2010)
7. Fogg, B.J.: Persuasive technology: Using computers to change what we think and do. Morgan Kaufmann (2002)
8. Petty, R.E., Cacioppo, J.T.: The elaboration likelihood model of persuasion. Advances in Experimental Social Psychology 19, 123–205 (1986)
9. Petty, R.E., Wegener, D.T.: The elaboration likelihood model: Current status and controversies. Dual-Process Theories in Social Psychology, 41–72 (1999)
10. Wiafe, I., Nakata, K., Gulliver, S.R.: Designing Persuasive Third Party Applications for Social Networking Services Based on the 3D-RAB Model. In: Park, J.J., Yang, L.T., Lee, C. (eds.) FutureTech 2011, Part II. CCIS, vol. 185, pp. 54–61. Springer, Heidelberg (2011)
11. Festiger, L.: A theory of cognitive dissonance. Row, Peterson, New York (1957)
12. Räisänen, T., Lehto, T., Oinas-Kukkonen, H.: Practical Findings from Applying the PSD Model for Evaluating Software Design Specifications. In: Ploug, T., Hasle, P., Oinas-Kukkonen, H. (eds.) PERSUASIVE 2010. LNCS, vol. 6137, pp. 185–192. Springer, Heidelberg (2010)
13. Fogg, B.J.: Captology: The study of computers as persuasive technologies. ACM, New York (1997)
14. Sterns, A.A., Mayhorn, C.B.: Persuasive Pillboxes: Improving Medication Adherence with Personal Digital Assistants. In: IJsselsteijn, W.A., de Kort, Y.A.W., Midden, C., Eggen, B., van den Hoven, E. (eds.) PERSUASIVE 2006. LNCS, vol. 3962, pp. 195–198. Springer, Heidelberg (2006)
15. Ajzen, I.: The theory of planned behavior. Organizational behavior and human decision processes 50(2), 179–211 (1991)
16. Radhakrishna, G., Saxena, A.: Application of cognitive dissonance theory to reduce dropouts in distance education system (2005)
17. Aronson, E.: Back to the future: Retrospective review of leon festinger's "a theory of cognitive dissonance". The American Journal of Psychology 110(1), 127–137 (1997)
18. Mathieson, K.: Predicting user intentions: Comparing the technology acceptance model with the theory of planned behavior. Information Systems Research 2(3), 173 (1991)
19. Fogg, B.J., Hreha, J.: Behavior Wizard: A Method for Matching Target Behaviors with Solutions. In: Ploug, T., Hasle, P., Oinas-Kukkonen, H. (eds.) PERSUASIVE 2010. LNCS, vol. 6137, pp. 117–131. Springer, Heidelberg (2010)

Towards a Data-Driven Approach to Intervention Design: A Predictive Path Model of Healthy Eating Determinants

Rita Orji, Regan L. Mandryk, and Julita Vassileva

Computer Science Department
University of Saskatchewan
Saskatoon, SK. S7N 5C9, Canada
rita.orji@usask.ca, {regan,jiv}@cs.usask.ca

Abstract. Dietary behavior and attitude play major roles in the worldwide prevalence of obesity, as weight is gained when energy intake exceeds energy expenditure. Although research has focused on designing technological interventions for healthy eating behavior, recent reviews have identified a gap in the knowledge base regarding the variables/determinants of healthy eating and the interactions between them. We developed a model of some determinants and their impact on healthy eating as a basis for designing technological interventions to promote healthy eating behavior within a target community. The main goal of this work is to understand how people adopt a healthy eating attitude, the variables influencing such attitudes, the interactions between these variables, and the degree of influence each variable exerts on healthy eating attitudes. We use fast food-related eating behavior as our case study. Our model shows that weight concern, nutrition knowledge, concern for diseases, social influence, and food choice motives predicts 65% of the variance in healthy eating attitudes, showing the suitability of the model for use in predicting healthy eating attitude. This result will inform decisions on the most effective persuasive strategy for designing interventions to promote healthy eating behavior.

Keywords: Dietary Behavior, persuasive intervention, predictive model, obesity, theories, determinants. healthy eating, fast food behavior.

1 Introduction

Obesity is a major health concern worldwide and the prevalence of overweight and obese individuals is especially alarming in developed countries. One of the few uncontroversial facts about obesity is that weight is gained when energy intake exceeds energy needs for a prolonged period. Eating behavior, therefore, becomes an important factor to consider in any interventions targeting obesity. The contributing effect of healthy diet in promoting general well-being cannot be overemphasized. Research has shown that good eating habits can reduce the risk of obesity, heart disease, and diabetes [1]. As a result, several interventions have focused on promoting

M. Bang and E.L. Ragnemalm (Eds.): PERSUASIVE 2012, LNCS 7284, pp. 203–214, 2012.

healthy eating attitudes [1, 2]. Most of these interventions are based on health behavior and health promotion theories adapted from various disciplines including psychology, sociology, consumer behavior, and marketing which are highly generalized [3]. Interventions that are informed by theories tend to be more successful than those based on intuition [3, 4]; however, designers of healthy eating interventions often face problems in adapting and applying these theories to the healthy eating behavior domain. This is because the theories are broad, abstract by nature, and lack specific content, therefore allowing room for subjective interpretations and applications. Recent studies have identified a need to determine the variables/determinants of a healthy eating attitude and the interactions between these variables to form the basis for healthy eating interventions [5]. Our study aims to fill this gap by understanding how people adopt a healthy eating attitude, the variables influencing attitude, the interactions between these variables, and the degree of influence each variable exerts on healthy eating attitude using fast food related behavior as case study. Fast food is defined as convenience foods obtained in self-service or 'take-away' eateries with minimal waiting. They are characterized by high energy dense, low in micronutrients and fiber, high in simple sugars and salt, majorly larger in portion size than conventional home-cooked or restaurant foods [6].

Persuasive technology researches have emphasized the need to study and understand the target user group to enable customization of interventions and to inform decisions on appropriate intervention strategies. We focus on a particular target group – university students. We conducted a mixed-methods study with 221 visitors in 10 fast food restaurants within the University of Saskatchewan (UofS) campus. The collection of primary survey data was followed by a 5-minute interview with 15 randomly selected participants. We employed Structural Equation Modeling (SEM) to explore the interaction between various determinants of healthy eating behavior from our survey. We employed SEM to exhaustively generate a predictive model of healthy eating behavior. The results show that weight concern, knowledge, concern for diseases, social influence, and food choice motive (listed in decreasing order of relevance) influence healthy eating attitude and predict 65% of the healthy eating attitude variance, showing the suitability of the model for use in predicting healthy eating attitude. The highest prediction of weight concern (50%) implies that concern for weight is the most significant influencer of healthy eating attitude in this target population and, therefore, should be emphasized by persuasive interventions targeting healthy eating.

Our work argues that behavioral predictor models that show the inter-relationship between various behavior determinants and their degree of influence on target behavior have a role to play in persuasive technology research. Our work complements the theoretical approach to intervention development by providing a practical way of testing theoretical variables on the target audience prior to intervention design.

2 Related Work

Persuasive technology is fundamentally about inducing behavior change using computers [7]. Developing effective persuasive interventions for behavior change

requires in-depth and practical understanding of human psychology [8]. The most effective persuasive interventions for behavior change usually occur when the intervention is behaviorally focused and theory driven [3, 8]. Therefore, persuasive intervention designers depend on health behavior and health promotion theories adapted from various disciplines to inform their work. Here, we present some relevant theories of health behavior.

Knowledge-Attitude-Behavior Model stresses the importance of knowledge as a prerequisite for intentional performance of health related behavior. An acquisition of new knowledge leads to changes in attitude, which in turn leads to an improved dietary behavior [9]. Research has distinguished between various types of knowledge based on degree of motivation. "Awareness knowledge" is a type of knowledge that captures people's attention, increases awareness, and enhances motivation, whereas "how-to" knowledge is the type people need when they are already motivated [10]. In other words, awareness knowledge enhances people's motivation to take action and instrumental knowledge is needed by people in order to act on their motivation [10].

Health Belief Model (HBM) is one of the first theories of health behavior that was developed to address problem behaviors that evoke health concerns [11]. The primary constructs of HBM include perceived susceptibility (perceived risk of contracting the health condition of concern), perceived severity (perception of the consequence of contracting the health condition of concern), perceived benefits (perception of the good things that could happen from undertaking specific behaviors), perceived barriers (perception of the difficulties and cost of performing behaviors), and cues to action (exposure to factors that prompt action).

While each of these traditional behavior theories is important, they complement each other rather than being mutually exclusive. Again, the impact of these theories could vary depending on the target group. Thus, a predictive model that determines the impact of various variables of the theories/models in a target audience is necessary.

3 Research Method

The data reported in this paper is part of a project aimed at designing persuasive intervention for promoting healthy eating that was approved by the UofS ethics board.

3.1 Research Approach and Measurement Instrument

This study employed both quantitative and qualitative methods of data collection. The quantitative component involved the collection of primary survey data from large numbers of participants with the intention of projecting the results on a wider population [12]. The qualitative part, which involved a post-survey 5-minute interview with 15 participants (7 females and 8 males between the ages 18 to 36) randomly selected from the study participants aimed to help elucidate the reasons behind the behaviors, and to clarify responses from the survey. This is in line with [13], which suggests that the most effective method of data collection is a combination of both quantitative and qualitative methodologies.

The questionnaire was developed after an extensive review of behavior change theories, persuasive interventions for healthy eating, and consumer behavior, and pilot

tested (n=10) for refinement. The survey instrument consists of questions assessing (1) participants' demography (gender, age group, education); (2) health concern (concern for weight and concern for diseases); (3) nutrition knowledge; (4) fast food motives; (5) healthy eating attitude; (6) frequency of purchase; and (7) social influence.

We adapted the 10 questions from Kahkonen [14] to measure the health concern variable. The scale has been validated by several studies [14, 15]. The health related concern measures the participants' degree of concern about food and health related issues using 5-point Likert value ranging from "1 = Not Concerned at all" to "5 = Very Concerned". Typical questions in this variable ask the participants to evaluate their degree of concern for "getting a lot of calories in food". The present study separated concern for health into two variables: Concern for Weight (WC) and Concern for Disease (DC). This is based on the factor loadings and the suggestion by Yu-Hua [16] that WC and DC might impact of health attitude/behavior differently.

Food choice motive measures several factors and their relative importance to the participants in making daily meal choices. The factors refer to health and non-health related food characteristics that might be taken into account when choosing what to eat. We adapted the 36 food choice motives questions developed by Steptoe et al. [17]. Some examples includes "It is convenient", "It is healthy", and "It is cheap". The present study allowed the participants to select from a list of factors that motivate their fast food consumption, which are mostly non-health related.

We measured attitude towards healthy eating using a 3-item scale adapted from Kearney et al. [18]. An example of a question in this category is "I make conscious effort to eat healthy". The participants state their level of agreement using a 5-point Likert scale, ranging from "1 = strongly agree" to "5 = strongly disagree".

To assess nutrition knowledge, we adapted the questions developed by Alexander [19] and used a 5-point Likert scale ranging from "1 = worst quality" to "5 = best quality". The questions were designed to solicit participants' knowledge about fast food meals by allowing them to rate the subjective nutrition quality of some selected fast food meals. A typical question is "Can you rate how nutritious and healthy you feel that French fries are?"

We included the social influence variable to determine the influence of others on purchase decisions. To do this, participants were presented with questions of the type: "Which of the following influence your decision to purchase fast food (you can select more than one): family, friends, colleagues, restaurant attendants, and self- decision". This particular question was deemed necessary because, although several researches have shown the important role that others play in motivating certain behaviors, the degree of social influence and its relationships with other variables are still unclear.

3.2 Research Participants

The participants consisted of 223 restaurant visitors sampled at selected fast food restaurants within the UofS campus. There were 221 usable responses. The participants were either students or employees of the UofS and the data were gathered over a period of 14 days in 2011. The only eligibility criterion required was that the participants were at least 18 years old at the time of data collection, in compliance with the study ethics approval and also to ensure that the participants were of legal age to make decisions independently (including decisions on what to eat). Gender was relatively evenly distributed; 45% (99) of the participants were female and 55% (122)

were male. The ages of the participants were sparsely distributed: 18-25 (153, 69%), 26-35 (55, 25%), 36-45 (10, 5%), and only 3 (1%) were over 45. Similarly, undergraduate students represent 80% of the university community [20], therefore, most of the participants in this study were high school graduates who are presently pursuing their first degree 104 (47%), 4 (2%) held a diploma degree, 52 (24%) held a bachelor's degree, 54 (24%) were master's degree holder, and 6(3%) were doctorate degree holders. Regarding the frequency of visit to fast food restaurants, more than 60% of the participants visited fast food restaurants at least 2 times in a week (every day: 7, 3%; 3 times a week: 62, 28%; once a week: 71, 32%; 3 times a month: 28, 13%; once a month: 39, 18%; never: 14, 6%). It is important to highlight that the exact frequency of visits might be greater as the interviewed participants revealed that fast food bought as take away or home delivered were not counted.

3.3 Data Validation

To ensure reliability and validity, we selected an analytical method that explicitly models the linear and quadratic effect (non-linear relationships) between the measured variables. We used the SmartPLS [21] Structural Equation Modeling (SEM) tool for simultaneous estimation of multiple equations.

Instrument Validation: To determine the validity of the survey instrument we performed Principal Component Analysis (PCA) using SPSS 19. Before conducting PCA, the Kaiser-Meyer-Olkin (KMO) and Bartlett sphericity tests were determined to measure the sample adequacy [22]. The KMO were all >0.7 and the result of Bartlett sphericity tests were significant at <0.001, thus, the data was suitable to conduct factor analysis [23]. The factor loadings and the corresponding factor scores (weights) the variable were generated. The factor loading resulted in removal of some questions and each factor has larger loading on its corresponding factor (≥0.7) than cross-loadings on other factors (≤0.4). Thus, these items could effectively reflect factors since they have good validity including convergent and discriminant validity [24].

Reliability of the Variables and Indicators: We examined the data for reliability using both SPSS and SmartPLS tool. To check for reliability, we used Cronbach's α, which efficiency ranges from 0 to 1 and can be used to describe the reliability of factors extracted. According to Peter [25], Cronbach's α should be ≥ 0.7, but for 2-3 indicator variables, a Cronbach's α ≥ 0.4 is acceptable. As shown in Table 1, column 4, the Cronbach's α of the variables satisfies these conditions (social influence contains 3 indicators, therefore, Cronbach's α is within the acceptable range of ≥0.4).

4 Results

After the validation of the data, we developed and tested the path model presented in Figure 1 using SEM in SmartPLS tool, which allows for simultaneous measurement (of indirect and direct influences of the variables) and structural models. In contrast to using SEM for hypotheses testing, our goal is to explore the interrelation/interaction between the various determinants of healthy eating and to generate a predictive model of healthy eating behavior. Therefore, we systematically examined the interactions and the impact of the five determinants (weight concern, concern for diseases, knowledge, choice motive, social influence) on each other and on healthy eating attitude. This enables us to exhaustively explore the significant of each determinants of healthy eating.

Table 1. Scale Reliabilities

Variables	AVE	Composite Reliability	Cronbach's Alpha	Redundancy (Q^2)
Threshold Value	**≥0.5**	**≥0.6**	**≥0.7 (2 to 3 indicators≥0.4)**	**≥0.0**
Choice Motivator	0.554	0.736	0.746	0.015
Concern for Disease	0.577	0.891	0.854	0.265
Healthy Eating Attitude	0.751	0.900	0.835	0.088
Knowledge	0.538	0.770	0.702	0.008
Social Influence	0.454	0.706	0.609	0.001
Weight Concern	0.523	0.839	0.753	0.001

4.1 Test of Proposed Path Model

Partial Least Square (PLS) model analysis essentially proceeds through two stages. The first stage deals with the reliability and discriminates validity analyses of the items and their associated independent variables in the outer model. The second stage estimates the relationships between the dependent variables in the inner model through bootstrapping procedures. Our analysis rigorously followed these two stages to confirm both discriminate and convergent validity and internal consistency. The model fit indices of the structural equation model is presented in Table 1. The square root of Average Variance Extraction (AVE) coefficients from the SmartPLS output is a key statistic at the first stage of the path analyses as it represents the variance extracted by the variable from its indicator items. As shown in Table 1, the AVE indices for all the variables are well above the theoretically ideal value of 0.5 with the exception of social influence, which is slightly below threshold (0.454). However, Cronbach's α values and the composite reliability that analyzes the strength of each indicator's correlation with their variables are higher than threshold values (see Table 1). Similarly, redundancy values are above "0" and the t-test values that measure the significance of the path coefficient are greater than 2.5 (above the recommended threshold of 1.96), as shown in Table 2.

The value of R^2 is shown in Table 2. The table shows that weight concern that determines 50% of variance is the highest predictor of healthy eating attitude followed by choice motivator 13%, and knowledge 2%. Overall, the model's independent variables predict 65% of the variance in healthy eating attitude showing the predictive relevance and the suitability of the model in predicting healthy eating attitude.

Again, to measure the shared variance between the variables and their measures, we evaluated the discriminate validity of the model. The discriminate validity further confirmed that the diagonal elements were significantly higher than the off diagonal values (i.e. correlation values) as shown in Table 3. Since all variables had diagonal elements (AVE) greater than the recommended value of 0.5 (except social influence) and also greater than the correlation values, the data demonstrates successful discriminate validation. Consequently, all calculated quality criteria indicated by SmartPLS are above the recommended threshold with the exception of social influence, which is slightly below the threshold of AVE but scores high in other fit indices and is therefore retained.

Table 2. T-ratio and R^2 for the Dependent Variables

Variables	R^2	t-test
CM	0.13	3.033
AT	0.37	5.500
KN	0.02	2.522
WC	0.50	12.446

Table 3. The AVE and Latent Variables Correlation Matrix

	CM	DC	AT	KN	SI	WC
CM	**0.554**					
DC	-0.208	**0.577**				
AT	-0.326	0.345	**0.751**			
KN	0.085	0.139	0.120	**0.538**		
SI	0.255	0.004	0.006	0.038	**0.454**	
WC	-0.121	0.384	0.396	0.049	0.084	**0.523**

CM = food choice motives, DC = concern for disease, AT = healthy eating attitude, N = nutrition knowledge, SI = social influence, WC = weight concern

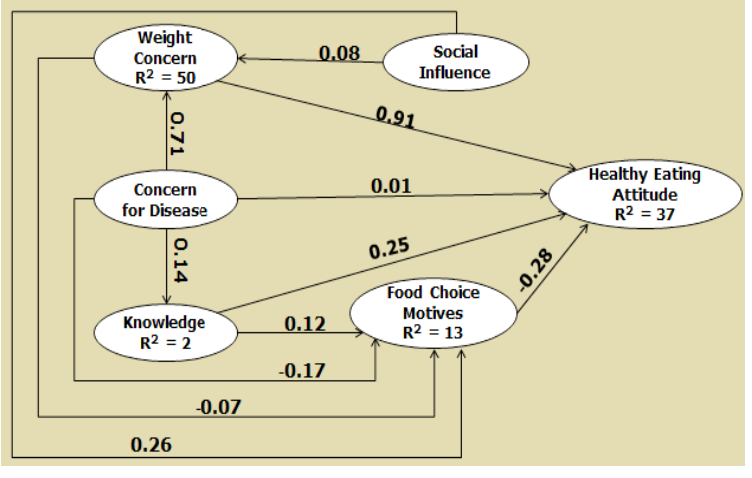


Fig. 1. A path model of healthy eating attitude, associated determinants' interrelations

4.2 Proposed Model

Figure 1 represents our proposed path model. In this model, concern for disease and social influence are the independent variables that influence the dependent variables (weight concern, food choice motive, and health attitude).

Interactions between the Model Variables

As shown in Figure 1, among all the variables, weight concern is most statistically and significantly associated with healthy eating attitude (with path coefficient β = 0.91). Concern for disease unfortunately, has no significant relationship with attitude; however, it exerts a strong positive effect on weight concern and knowledge. Similarly, weight concern and concern for disease reduces the impact of food choice motives on attitude. Knowledge on the other hand positively influences both healthy eating attitude and food choice motive. Contrary to expectations, social influence, interestingly, have no direct effect on healthy eating attitude and does not significantly interact with any other variable apart from food choice motive, for which it exerts the most significant effect. Again, the model shows that food choice motive is the single negative determinant/influencer of healthy eating attitude. The path values (interactions) presented in this model are statistically significant at $p \leq 0.01$.

5 Discussion

The main goal of this study is to use fast food related behavior as a case study, to understand how people adopt a healthy eating attitude, the variables influencing attitude, the interactions between the variables, and the degree of influence each variable exerts on healthy eating attitude. Our model serves as the basis on which to make decisions on the most effective persuasive strategy and approach to employ in the design and development of interventions targeting healthy eating behavior (especially when eating in restaurants). From thorough literature review, this study is the first to exhaustively examine the combined interaction of the variables weight concern, concern for disease, knowledge, food choice motive, social influence, and health eating attitude, and their influence on attitude towards healthy eating. The study also examined the mediating role played by food choice motive.

Health Concern (Weight Concern and Concern for Disease) and Eating Attitude: Surprisingly, although, participants expressed health concerns from consuming fast food in general, the study shows that concern for disease is the least important motivator (as shown in Figure 1) while concern for weight stands out as the most significant motivator of healthy eating attitude. This contradicts the finding by Yu-Hua [13] that weight concern is the least important motivator in making food-related decisions. One possible explanation might be the age group of this study. The majority of our study participants are young, falling within the age group 18-25 (69%) and 26-35 (25%). Most people within this age range believe that having an attractive image or socially desirable traits is associated with having a good personality that is the "physical attractiveness stereotype" [26]. Similarly, according to Kai-Yang [27], people believe that good body image is linked to life of happiness, success, and social acceptance the same way that fatness is associated with laziness, stupidity, and chaos. Our study suggests that physical self-presentation and social physique anxiety are keys to healthy behavior motivation among younger adults. Another likely explanation as noted by Yu-Hua [16], is that "food intake plays only a partial role in the onset of various diseases, while it is usually the only avenue for obtaining calories", hence individuals who exhibit concern of gaining weight from food are more inclined to a healthy eating attitude than those who exhibit concern for developing diseases. It is however important to state that obsession for weight as opposed to diseases control might not be the case for an older population, who are likely to be more sensitive to their general health as they are more prone to acquiring diseases [28].

It is also interesting to highlight the influence of both weight concern and concern for disease on food choice motive. The model shows that they negatively influence food choice motive. This means that individuals motivated to eat healthily due to their concern for health (concern for weight and disease) are less affected by non-health related food choice motives (e.g. convenience, tastes good) and can easily overcome the barriers presented by these factors towards healthy eating attitude. Similarly, although, concern for disease produces no significant direct influence on attitude, it strengthens weight concern and knowledge. Therefore, people who are already concerned about their weight are more likely to also care about disease than people who are not. Notably, concern for disease increases the quest for health-related

knowledge; however, it might not be an advisable key motivator of behavior change for this target group.

Knowledge and Healthy Eating Attitude: As expected, knowledge positively influences attitude. Knowledge is one factor that has generated contradicting views and results from researchers in health intervention. Brown et al. [29] suggest that a high level of nutrition awareness does not reflect in participants' food preferences while Strychar et al. [30] and Leme et al. [31] show a positive link between knowledge and healthy eating. Leme et al. however, stressed that knowledge about healthy eating is not the only determinant of healthy food choice. Therefore, traditional nutritional interventions focusing on healthy eating should be modified to aim at behavioral change. These mixed findings are not surprising considering that none of the researchers have examined the impact of knowledge in line with an individual's health goal. According to Fogg [4], persuasive interventions are more effective when they align with an individual's health goals. One of the findings of our study (although beyond the scope of this paper) shows that individuals have varying health goals. Most knowledge-based healthy eating interventions can be described as arbitrary to most people as they are based on one-goal-fits-all approach, and are therefore, hardly useful. A tailored knowledge will likely increase healthy attitude as it will act as a facilitator in line with an individual's health goals. As Carrillo et al. [32] noted, most food with relative health value often exhibits low consumption frequency due to low knowledge about their health benefit.

Social Influence and Healthy Eating Attitude: Interestingly, the model suggests that social influence has no direct impact on healthy eating attitude but produces an indirect positive effect via food choice motives and weight concern. From literature, social influence is one of the most widely employed persuasive strategies. However, little is known about the interaction between this variable and other behavior change motivators. Our model reveals that social influence is not likely to act as a primary motivator of change; however, it can increase the effectiveness of other motivators. This can find interpretation in friendship patterns and grouping among university students (young adults) who tend to be attracted to people of similar inclinations, therefore, leaving little or no room for intra-group influences. From our post survey interview, 4 participants in the study group who are reportedly overweight and eat fast food at least 3 times a week confirmed that most of their friends exhibit similar behaviors and therefore, it is a normal behavior within their circle of friends. This presents an interesting finding that suggests that inter-group intervention might be necessary to produce significant healthy eating change among this circle of friends. The finding also suggests that self-presentation and social group anxiety might be strongly mediated by group influence as people tend to care and seek the approval of similar others (in-group acceptance) more than they do for any external human influencers [1]. This has been confirmed by a survey that reported that 79% of the students eat lunch with friends [33].

The Mediating Role of Choice Motives: The model indicates that food choice motive (e.g., convenience, time saving) negatively influences healthy eating attitude. However, individuals who are concerned about their weight and diseases will have a lesser effect of food choice motive and will make a conscious effort to eat healthily. Hence, weight concern positively influences healthy eating attitudes both directly and

indirectly via food choice motive for health [16]. Social influence is the only variable that produces a significant positive effect on food choice motive. The explanation follows closely from the results of social influence. Friends eat together as social bonding, and fast food restaurants are convenient, fast, and inexpensive places to do this. Notably, food choice motive reduces healthy eating tendencies. This is not surprising as most motives for fast food consumption are non-health related (e.g., fast, cheap, fun). Overall, our results suggest that persuasive researchers targeting healthy fast food consumption should plan to deal with the inhibiting effect of food choice motive on healthy eating attitude.

6 Implication

Recent persuasive technology research has emphasized the need to study and understand the focus group of interest to enable customization of interventions and decision on appropriate intervention strategy. Similarly, it is a common practice for persuasive interventions to employ multiple or a combination of strategies in a design (to increase the persuasiveness of their applications) with the hope that at least one out of the many strategies will suit the target audience. This approach, however, unnecessarily increases the complexity of the intervention, and therefore, contradicts the need to make technological interventions usable (useful and simple). It also makes it hard to pinpoint the most effective strategy that actually motivated the behavior change. To solve these problems, there is a need to understand the target audience, the important variables, and the interactions between these variables. This will not only inform the design decision on effective strategy but will make it easier to imitate any successful strategy or to pinpoint the strategy that produced the highest degree of influence on the target behavior.

Our findings have both theoretical and practical implications. On the theoretical front, our study shows how models can be used to predict behavior. Specifically, they highlight the separate role of concern for disease and weight, nutrition knowledge, and social influence on healthy eating attitude. The results strongly support the assertion that food choice motives have independent and mediating effects on healthy eating attitudes. We hope that this study spurs research into examining the role of these variables and other variables of interest on other health behaviors, such as exercise. Our work argues for the use of models as the basis for behavior intervention design.

From a practical point of view, our results shed light on important variables that individuals consider when forming healthy eating attitudes. Among all the variables, concern for weight exhibits the strongest direct influence on healthy eating attitude. Thus, persuasive interventions targeting healthy eating in younger adults should emphasize the relationship with weight rather than diseases. Furthermore, the model suggests that knowledge is an important factor in the design of healthy eating persuasive interventions. This is an area that has been largely ignored by persuasive researchers; however, our results show that the addition of a knowledge component has both direct and indirect effects that help healthy eating attitude formation. This means that integrating healthy eating education in persuasive interventions may be helpful. Similarly, the impact of social influence need not be ignored; although social influence has no direct influence on healthy eating attitude, it increases the effect of weight concern on healthy eating attitude.

Thus, to achieve better result, persuasive interventions should use social influence along with any strategy emphasizing weight concern.

This research study has some limitations that must be acknowledged. One practical limitation for generalizability stems from younger participants from the university community, the majority of whom are students. We are careful not to generalize for a population older than 35.

7 Conclusion

In general, we argue that behavioral predictor's models which show the inter-relationship between various behavior determinants and their degree of influence on target behavior have a role to play in persuasive technology research. The findings support the use of models as a basis for the study of eating behavior and the influence of various healthy eating determinants. The model can also serve as a foundation to inform an effective intervention. In summary, the results show the need to emphasize weight concern when designing persuasive interventions especially those that target young adults. Also, worth considering is the impact of knowledge and the mediating role that food choice motive play. Overall, this result suggests that persuasive researchers targeting healthy fast food consumption should effectively plan to deal with the inhibiting and mediating effect of their food choice motive on healthy eating attitude. The finding also shows that our model's variables predicted a sizable percentage of variance (65%) in healthy eating attitude, indicating the appropriateness of the model in predicting healthy eating attitude. The work contributes to the persuasive interventions literature by illuminating how various variables influence healthy eating attitude and the relationship between these variables. Future research may consider modifying the model to improve its predictability, this include adding other important determinants and mediators.

Acknowledgements. The first author of this paper is being sponsored by Vanier Canada Graduate Scholarship. Many thanks to Richard K. Lomotey for his assistance with the data collection and to the reviewers for their insightful comments.

References

1. Wansink, B.: Mindless Eating: Why We Eat More Than We Think. Bantam Books (2006)
2. Winett, et al.: Nutrition for a lifetime system: a multimedia system for altering supermarket shoppers' purchases to meet nutritional guidelines. Compt. Hum. Beh., 371–392 (1997)
3. Glanz, K., et al.: Theory at a glance: A guide for health promotion practice. U.S.A Dept of Health and Human Services. Public Health Service. National Institutes of Health (1995)
4. Fogg, B.J.: Mobile Persuasion: 20 Perspectives on the Future of Behavior Change (2009)
5. Government of Canada. Promotion and Support of Healthy Eating: An Initial Overview of Knowledge Gaps and Research Needs, Summary Report, Health Canada (2003)
6. Feeley, et al.: Fast-food consumption among 17-year-olds in the Birth to Twenty cohort. S. Afr. J. Clin. Nutr. 22(3), 118–123 (2009)
7. Fogg, B.J.: Persuasive Technology: Using Computers to Change What We Think and Do (2003)
8. Fogg, B.J.: A behavior model for persuasive design. In: Persuasive 2009, pp. 1–7 (2009)

9. Freeland-Graves, et al.: Position of the American Dietetic Association: Total Diet approach to Communicating Food and Nutrition Information. J. Am. Diet. Assoc., 100–108 (2002)
10. Nweze, E.N.: Community nutrition: planning health promotion and disease prevention. Jones & Bartlett Learning (2009)
11. Rosenstock: Why people use health services. Milbank Memorial Fund Qurt., 94–127 (1966)
12. Tustin, D.G., et al.: Marketing research in practice. UNISA, Pretoria (2005)
13. Datta, L.: Paradigm wars: A basis for peaceful coexistence and beyond. In: Reichardt, C.S., Rallis, S.F. (eds.) The Qualitative-Quantitative Debate: New Perspectives (1994)
14. Kähkönen, P., Tuorila, H., Rita, H.: How information enhances acceptability of a low-fat spread. Food Quality and Preference 7(2), 87–94 (1996)
15. Bower, J.A., et al.: Effect of liking, information and consumer characteristics on purchase intention and willingness to pay more for a fat spread with a proven health benefit. Food Quality and Preference 14, 65–74 (2003)
16. Yu-Hua, C.S.: Health concern, food choice motives, and attitudes toward healthy eating: The mediating role of food choice motives. Appetite 51, 42–49 (2008)
17. Steptoe, et al.: Development of a measure of the motives underlying the selection of food: The Food Choice Questionnaire. Appetite 25(3), 267–284 (1995)
18. Kearney, et al.: Attitudes towards and beliefs about nutrition and health among a random sample of adults in the Republic of Ireland and Northern Ireland. PHN, 1117–1126 (2001)
19. Alexander, J.M.: Dieting strategies of college students: a study of behaviors, attitudes and goals, and the influence of restraint and gender. M.S. Thesis, Rutgers Univer (1992)
20. http://www.usask.ca/isa/statistics/students/headcount-demographics.php
21. http://www.smartpls.de/ (accessed December, 2011)
22. Kaiser, H.: A Second Generation Little Jiffy. Psycometrika 35, 401–415 (1970)
23. Freeland-Graves, et al.: Position of the American Dietetic Association: Total Diet pproach to Communicating Food and Nutrition Information. J. Am. Diet. Assoc., 100–108 (2002)
24. Gefen, et al.: Structural Equation Modeling and Regression:Guidelines for Research Practice. Communications of the Association for Information Systems, 1–70 (2000)
25. Peter, S.: Kundenbindunq als Marketingziel: Identifikation und Analyse zentraler Determinanten, Gabler, Wiesbaden (1997)
26. Dion, K.: Stereotyping based on physical attractiveness: Issues and conceptual perspectives. In: Herman, C.P., Higgins, M.P. (eds.) Physical Appearance, stigma (1986)
27. Kai-Yang: The Relationship between Physical Self-Presentation and Social Physique (2002)
28. CDC's Disaster Planning Goal: Protect Vulnerable Older Adults (accessed December 2011)
29. Brown, K., McIlveen, H., Strugnell, C.: Nutritional awareness and food preferences of young consumers. Nutrition & Food Science 30(5), 230–235 (2000)
30. Strychar, I.M., et al.: Changes in knowledge and food behavior following a screening program held in a supermarket. Canadian J. of Public Health 84(6), 382–388 (1993)
31. Leme, A.C.B.: Association of Brazilian Adolescents with Healthy Eating: Knowledge, Perceptions and Food Choices. Food and Nutrition Sciences, 1036–1042 (2011)
32. Carrillo, E., et al.: Main Factors Underlying Consumers Food Choice: A First Step For The Understanding Of Attitudes Toward Â Healthy Eating. J. of Sensory Stud., 85–95 (2011)
33. http://www.kvhigh.com/PSSC/healthyeating.htm (accessed December, 2011)

Passengers' Safety in Aircraft Evacuations: Employing Serious Games to Educate and Persuade

Luca Chittaro

Human-Computer Interaction Lab
University of Udine
via delle Scienze 206
33100 Udine, Italy
http://hcilab.uniud.it

Abstract. The field of persuasive technology has only recently started to investigate how virtual experiences of risk can be used to change people's attitudes and behaviors with respect to personal safety. In this paper, we aim at advancing the investigation in different directions. First, we extend the study to self-efficacy, which has been shown to be a predictor of future performance as well as an important factor for persuasion attempts which show negative consequences on people's health. Second, we increase the interactivity of the virtual experience, by designing and implementing a full serious game, in which the user can acquire knowledge about several aspects of her personal safety, and we investigate also effects of the virtual experience on user's knowledge. Third, we focus on an important problem to which serious games and persuasive technology have never been applied before, i.e. educating passengers about personal safety in aircraft evacuations. The experiment presented in the paper shows how just playing the serious game for a few minutes results in significant increases in user's knowledge and self-efficacy.

Keywords: virtual reality, personal safety, serious games, simulated risk experiences, self-efficacy, risk perception, air passengers, aircraft evacuation.

1 Introduction

The field of persuasive technology has only recently started to investigate how virtual simulations of risk experiences can be used to change people' attitudes or behaviors with respect to personal safety [8,21,30]. In general, simulation can persuade people to change by enabling them to observe immediately the link between cause and effect [13]. Moreover, if a virtual risk experience is interactive, it can exploit operant conditioning, by allowing the user to choose his/her behaviors in the virtual experience and providing immediate feedback by showing the positive consequences of recommended behaviors and the negative consequences of dangerous behaviors. Those consequences can be simulated in vivid and memorable ways through visual and auditory stimuli: in this way, the simulation can include affective aspects, which contribute significantly to determine risk perception [28].

M. Bang and E.L. Ragnemalm (Eds.): PERSUASIVE 2012, LNCS 7284, pp. 215–226, 2012.

The experiments on persuasive virtual experiences of risk carried out so far have focused on measuring their effects on attitudes towards global climate change [21], information search and coping intentions towards flood risks [30], perception of fire risks in buildings [8].

In this paper, we aim at advancing the investigation of virtual risk experiences in different directions. First, we extend the study of their effects to self-efficacy, which is in general an important predictor of future performance [3]. In the specific case of virtual risk experiences, self-efficacy is important also for an additional reason. These experiences can be threatening, and protection motivation theory (PMT) [25] points out how the consideration of self-efficacy is necessary for the success of persuasion attempts which inform people about the negative consequences of given actions on their health. Indeed, if the persuasive attempt threatens the individual, but does not make her feel capable of performing the recommended actions, then PMT predicts that, instead of being persuaded, she will try to reduce the negative emotions induced by threat, e.g. through risk denial and defensive reactions.

Second, we increase the interactivity of the simulated experience with respect to previous studies, by designing and implementing a full serious game (i.e. a videogame to further training and education objectives [31]), in which the user can experiment the whole range of possible (right or wrong) actions typical of the considered risk experience, making progress towards game level completion when she chooses the right ones. In this way, the user can acquire knowledge about several aspects of her personal safety, and we thus investigate also effects of the virtual experience on user's knowledge.

Third, we focus on educating passengers about their personal safety in aircraft evacuations, an important problem to which serious games and persuasive technology have never been applied before.

The paper is organized as follows. In Section 2, we discuss in detail the need for more persuasive approaches to passengers' education in aviation safety, by examining the limitations of the currently employed solutions and motivating our proposal. Section 3 illustrates the importance of the self-efficacy construct in the domain of safety, and describes the persuasive goal and target behavior we consider. In Section 4, we present the serious game we have implemented. Section 5 and 6 respectively present the experimental evaluation and the obtained results. Section 7 concludes the paper and introduces future work.

2 Aviation Safety and Passenger Behavior: A Need for More Persuasive Approaches

Fast and safe evacuation of aircrafts during emergencies is a fundamental aspect of aviation safety. The need for high evacuation efficiency is explained by the fact that the aircraft cabin becomes an unsurvivable environment in about two minutes since fire erupts [22].

Unfortunately, incident and accident reports describe a wide range of inappropriate behaviors by air passengers during emergency situations and aircraft evacuations that

jeopardize their and others' survival. For example, in a large safety study conducted by the NTSB [23] which interviewed 457 passengers who have been involved in emergency evacuations, a large number of them (about 50%) admitted to try bringing their luggage with them during the evacuation, thus slowing down the process. Another typical passengers' error is to engage in competitive behaviors with other passengers such as pushing or trying to jump over rows of seats, which can make the evacuation chaotic and considerably slow it down. Additional issues concern lack of knowledge and ability, e.g. not going for the closest emergency exit, not being able of opening doors, moving in smoke instead of crawling below it, trying to carefully sit or to walk on the emergency slide instead of jumping and then sliding.

The primary purpose of aviation safety education is to provide airline passengers with accurate cabin safety knowledge and to cultivate positive passenger attitudes to appropriately affect passenger behavior when an emergency occurs. As pointed out in the study conducted in [7], the level of aviation safety education an airline passenger has does affect her knowledge, attitudes and behaviors. Safety awareness can lead passengers to efficient behaviors and being responsible for their own safety; therefore, improving passenger safety education will increase the probability of their survival in an emergency [22,26].

Education and training is also important because stress and negative affect in a real emergency, combined with lack of knowledge about suitable behaviors, produces in some passengers a "cognitive paralysis" phenomenon, where people do not take any action at all, leading to fatalities in otherwise survivable conditions [18,19]. Therefore, clearly knowing in advance proper safety behavior is a personal protection strategy that would allow each passenger to significantly increase her chance of survival by reducing evacuation time as well as preventing common fatal errors that passengers make due to lack of knowledge. Preparedness also contributes to reduce stress and fear caused by emergency situations, and air passengers need to know the most common potentially hazardous circumstances [12].

Current approaches to passenger education are based on the safety card and flight attendant presentation to which passengers are exposed after they have boarded the aircraft. Unfortunately, these approaches suffer from serious limitations [9]: passenger safety briefings and cards vary greatly; passenger attention to them is poor at best; comprehension of safety cards by passengers is below acceptable limits; studies have shown that typical passengers - even those who report that they pay attention to passenger safety briefings and cards - have little personal knowledge and understanding of the information they have been given to improve their chances of survival. As a result, one of the major reasons for deaths and injuries which could be preventable is that passengers lack preparedness.

The inadequacy of safety cards and briefings is confirmed by passengers who have been involved in real emergencies. For example, Chang and Yang [6] studied the emergency evacuation experiences of 110 passengers involved in a recent serious accident (China Airlines Flight CI-120) to examine deficiencies in passenger safety education: only 14% (respectively 16%) of the passengers found the safety briefing (respectively safety cards) to be useful with respect to their actual need of evacuating the burning aircraft. The majority of passengers said that the received safety

information is not sufficient for dealing with emergency escape and they did not feel to have been clearly instructed.

To increase the probability that air travelers will survive in emergencies, substantially improved safety and survival information needs to be implemented and made available through well-constructed passenger education [10]. Different authors agree that airlines should make preflight safety information more appealing and more comprehensible. Cosper and McLean [10] recommend to consider the development of state-of-the-art methods using "creative technologies" for passenger education such as interactive CD-ROMs that could be passed out at airports, air shows, and public events. Chang and Yang [6] suggest that civil aviation authorities should build a safety education exhibit at all airports with safety equipment and emergency use procedures to give passengers an opportunity to use and understand them.

This paper proposes to use serious games as a tool to develop personal safety skills for the following reasons. Compared to safety education exhibits in airports, serious games could be a less costly solution that would also allow passengers to live the simulations discreetly at their homes whenever and how many times they want. Compared to safety cards and video CD-ROMs, training passengers through a game would allow to make safety education materials more attractive and to simulate aircraft emergencies in a much more thorough and realistic way. Indeed, a serious game can immerse its user in aircraft emergency scenarios, where the goal of the game is to survive the emergency and player's survival is strictly dependent on choosing the right actions and taking as less time as possible to complete the evacuation, while staying as far away as possible from danger. To succeed and progress in the game, users would need to improve their decision making in aircraft emergencies, learning to avoid common passengers' errors.

Since people could be willing to devote more attention to a game, and the game could be played at home, the serious game solution could increase exposure time to personal safety content, and promote repetitive rehearsal of safety procedures, which improves retention of knowledge. Moreover, the game could take place in high-fidelity 3D reconstructions of actual airliners. This would allow not only to learn general knowledge that applies to any aircraft emergency (e.g., avoiding smoke, leaving luggage on the plane,…), but would also allow people to familiarize with the different escape routes, seat configurations, location and operation of emergency doors and slides available in the actual aircraft type they are going to fly with.

3 Self-efficacy and Personal Safety

Self-efficacy can be defined as the person's belief in his or her ability to perform a specific behavior [13]. According to Bandura's [2,3,4] Social Cognitive Theory, this belief significantly determines performance outcomes, and different people with similar skills may perform differently depending on variations in their self-efficacy. Research on self-efficacy has shown that the conviction that one can successfully execute the behavior required has a positive effect on performance [3,4].

Positive associations between safety training, self-efficacy and attitudes toward safety have been found in the literature, confirming the importance of the self-efficacy construct also in the field of safety (for a summary, see [15,16]). Gaining experience in performing the given behavior is a major factor that contributes to increase self-efficacy [3]. In this sense, the simulations that people can live with a serious game allow them to actually succeed in applying safety knowledge to a virtual life experience, instead of passively listening to traditional safety messages.

Increasing self-efficacy is particularly important in the domain of air passengers' personal safety. Indeed, while in other types of emergencies such as fires in buildings people tend to downplay the severity of the risk to their safety and overestimate their ability to move in the dangerous environment [24], people tend instead to be fearful of even normal flying conditions: estimates of the percentage of population which suffers from fear of flying (which can range from continuous apprehension about flying to severe phobia that can prevent flying) reach up to 40% [29]. Moreover, people have a pessimistic and fatalistic attitude towards aircraft accidents, mistakenly believing that there is little hope of survival. In fact, statistics show otherwise: a survey of commercial jet airplanes accidents [5] indicates that the majority of aircraft accidents is survivable. Another reason why many passengers do not pay attention to safety information is that they tend to shift the responsibility and capability of their safety to the cabin crew [22]. This way of thinking is dangerous because workload and the time constraints of the evacuation makes it impossible for the crew to provide individual assistance to every passenger. Besides, members of the crew could be injured or incapacitated, and this would require passengers to take an even more active role to survive.

Therefore, while increasing perception of risk severity (by also appealing to fear [8]) is a priority for fires in buildings, in the case of aircraft accidents people need instead to be persuaded about their ability to act to increase their likelihood of surviving the emergency.

Of the three possible reasons that prevent desired target behavior highlighted by Fogg's FBM model [14], lack of ability is the one that applies most to the case of response to aircraft emergencies. Indeed, lack of motivation towards proper behavior is unlikely (most people want to survive the emergency), and the lack of a proper trigger can be excluded (clear visual, auditory, olfactory and/or haptic cues present themselves to trigger the behavior). The problem is to persuade passengers that they are capable of acting properly and choosing the right behavior in response to the trigger. A serious game could be an ideal tool to this purpose, presenting the player with the effects of her wrong or right choices in a memorable way.

4 The Serious Game

The serious game we have created allows users to realistically experience aircraft evacuations scenarios of different severity and complexity and try for themselves the effects of taking the different possible actions. We have built an accurate 3D model of the cabin of an Airbus 320 [1], one of the most used aircraft types in service. The

simulation includes realistic sounds (e.g., the message that can be heard on the plane to inform passengers they have to brace, some shouts from other passengers,…) and visuals (e.g. smoke and fire effects). Since the main goals of the persuasive game concern increasing knowledge and self-efficacy of passengers, and not risk perception of aircraft emergencies (which as we have seen before people already perceive as serious), we chose not to make the virtual experience too emotionally intense by omitting the portrayal of the character's harm and distress. For example, if the character stands inside toxic smoke, the game level stops and the player is informed about the consequences only textually: she does not see the details or hear the sounds of the character death by suffocation as was instead done in [8], which wanted to increase perception of risk severity about fires in buildings.

The level played by the participant in the study described in this paper concerns an emergency landing in which the player has to face and can learn about several behaviors that greatly affect her personal safety, that is: (i) locating the nearest exits before the emergency occurs, (ii) maintaining the brace position during the emergency landing until the plane comes to a stop, (iii) avoiding taking luggage with oneself during the evacuation, (iv) reaching for the nearest exit, (v) locating an alternative exit in the presence of blocked exits or exits which have been reached by fire, (vi) avoiding competitive actions (pushing or fighting with other passengers, jumping on rows of seats), (vii) crawling below smoke, (vii) jumping on the slides instead of trying to slowly sit on them or to stand on them.

The player can succeed or fail at each of these steps. If she chooses the correct action, then she progresses in the game. When a wrong action is chosen, after learning about its consequences through a short textual description, the player is brought back to the part of the game level in which she took the wrong decision and she has the opportunity to restart from that step, instead of restarting from scratch. As pointed out by [17], this way of organizing a virtual experience can be seen as an application of the reduction and tunneling strategies [13] used in persuasive technology.

To make the game accessible to a wide audience, we did not assume specific experience with videogames or the availability of game input controls: at each moment during gameplay, all actions can be taken by using the four arrow keys and the Ctrl key which is close to them. If the action involves navigation, the four arrow keys control movement, and we have constrained the paths of the virtual character in such a way to prevent the typical situations in which inexperienced players get stuck on 3D objects in the environment (the purpose of the game is indeed to focus on personal safety content and procedures, not on learning to fine control video game characters as in entertainment games). For example, if the character is seated in the aircraft with the belts unbuckled and the player presses the up key, the character will stand up; then, if the player presses the right or left keys, the character will move in a natural way towards the aisle or the window, provided that there are no other passengers which are blocking the path. For additional actions (such as taking luggage, crouching in smoke, trying to push other passengers,…) a semitransparent legend appears in the lowest line of the screen to inform the player about currently available actions and their mappings with keys.

Figures 1 and 2 show examples of respectively the player's character correctly assuming the brace position and a situation in which the player is facing a double threat posed by unusable wing exits and presence of toxic smoke.

Fig. 1. Assuming the brace position

Fig. 2. Unusable wing exits and toxic smoke threat

5 Method

5.1 Participants, Design and Measures

We recruited 26 participants (19 male, 7 female) through personal contact. Participants were volunteer university students who received no compensation and their mean age was 23.85 (SD=2.51).

Video game use was assessed by asking participants to rate their frequency of use of video games on a 5-point Likert scale (1=never, 5=several hours a day) and their liking of video games on another scale (1=not at all, 5=a lot). Mean frequency of use was 2.50 (SD=1.33) and mean liking was 3.35 (SD=1.09).

We also asked participants about how often they travel by air on a 5-point Likert scale ranging from never to frequent flyer. The mean was 2.38 (SD=1.17). Liking of the flight experience on another scale (1=not at all, 5=a lot) resulted in a mean of 3.04 (SD=1.25).

To measure participants' knowledge about the safety aspects which were dealt with by the game level employed in the evaluation, we used a safety questionnaire with 6 multiple-choice items about (i) proper behavior before an emergency landing, (ii) what to do when the aircraft comes to a stop in the emergency landing, (iii) what to do in case of smoke in the cabin, (iv) usage of slides, (v) behavior in case of exit blocked, (vi) available exits on the aircraft. Four possible answers were presented for each question: three corresponded to typical passengers' errors and one was correct.

To measure subjects' self-efficacy, we designed a 7-item questionnaire, by (i) taking items from well-known self-efficacy questionnaires such as the General Self-Efficacy (GSE) scale [27] and adapting them to our domain description, e.g. "I am confident that I could deal with an emergency evacuation of an aircraft", and (ii) following the recommendation on rigorous theory-based semantic structure for specific behaviors proposed by [20], which leads to build items such as: "I would be able to deal with an emergency evacuation of an aircraft even if there was smoke in the cabin" or "I would be able to deal with an emergency evacuation of an aircraft even if some exits were blocked". Each item was rated by participants on a 5-point Likert scale (1=not at all, 5=very), so the questionnaire assigns a score ranging from 5 to 35 to measure participant's self-efficacy with respect to aircraft evacuation. We assessed internal reliability of the designed questionnaire with Cronbach's alpha test which indicated very high reliability (α =.94).

We also measured risk perception by using the 6 questions employed by [11]: vulnerability to risk was assessed by having respondents rate their vulnerability on 3 items (e.g., "how high do you believe your risk of being involved in an aircraft evacuation is?") and severity of risk on the 3 other items (e.g., "how harmful would the consequences of an aircraft evacuation be?"). Ratings were given on a 5-point Likert scale (1=not at all, 5=very), so the score for each of the two measures ranged from 5 to 15.

Considering the design choices on which the game is based (see Section 4), we hypothesized that playing the game should increase participants' self-efficacy as well as increase their level of knowledge about aircraft evacuations, while it should not heighten their risk perception. As a measure of the change in self-efficacy, we took the difference between its measure taken after and before the game experience, and we proceeded in the same way with risk perception. As a measure of the change in knowledge, we took the difference in the number of wrong answers to the safety questionnaire, taken after and before the game experience.

5.2 Procedure

Subjects were welcomed in the lab and told they were going to try a video game that illustrates procedures of aircraft evacuation. They were clearly informed that they could decide to refrain from continuing the experiment at any time without the need for providing a reason to the experimenters.

First, subjects filled the demographic, knowledge, self-efficacy and risk perception questionnaires. Then, the experimenter instructed them about the simple game controls and checked if they had understood their usage before allowing them to start playing the game.

Participants played the game on a 30 inches LCD monitor with stereo speakers. The opening screen of the level showed an external view of the aircraft flying low over terrain, with a brief text that introduced the scenario in which the captain had just informed the passengers that he was going to attempt an emergency landing. Then, the viewpoint moved inside the cabin and the game action started with the words "Brace! Brace!" aired on the aircraft public address system. We let participants play until they successfully completed the level, which took between 2 and 3 minutes. After participants completed the level, they answered the knowledge, self-efficacy and risk perception questionnaires for the second time.

6 Results

The means for number of wrong answers and self-efficacy score measured before and after the experience are shown in Figure 3. Differences were analyzed with a non-parametric Wilcoxon test and confirm our hypotheses: after playing the game, there was a statistically significant ($Z=-4.18$, $p<0.001$) decrease in the number of wrong answers which was more than halved, moving from 2.85 (SD=.93) to 1.38 (SD=.98), and a statistically significant increase in self-efficacy ($Z=-4.27$, $p<0.001$) which rose by 27%, moving from 17.38 (SD=5.91) to 22.27 (SD=5.27).

The means for risk vulnerability and severity measures before and after the experience are shown in Figure 4. Differences were analyzed with a non-parametric Wilcoxon test and confirmed our hypothesis that risk perception was not going to be heightened: after playing the game, there was no significant change in vulnerability perceptions. However, it is interesting to note that there was a statistically significant ($Z=-2.89$, $p<0.01$), although relatively small decrease in severity, which declined by about 10%, from 3.61 (SD=.85) to 3.23 (SD=.82). This change might be related to the fact that the simulation was deliberately designed not to be too intense emotionally, but also to the increase in self-efficacy. There was indeed a statistically significant negative correlation between the level of self-efficacy and perception of risk severity, before ($r(26)=-.57$, $p<.01$) as well as after ($r(26)= -.46$, $p=.018$) playing the game.

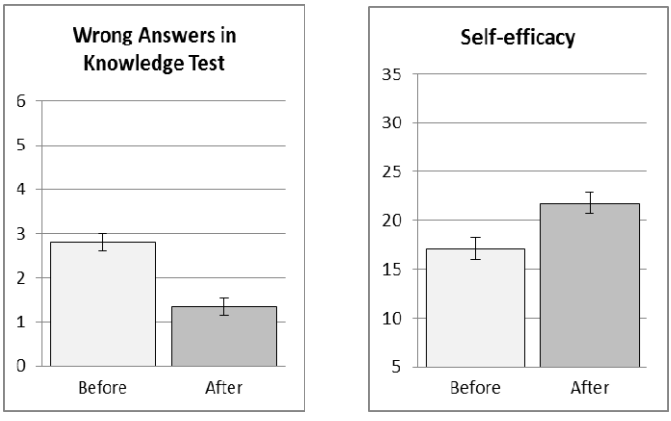


Fig. 3. Mean number of wrong answers and self-efficacy, before and after the experience. Capped vertical bars denote ±1 SE.

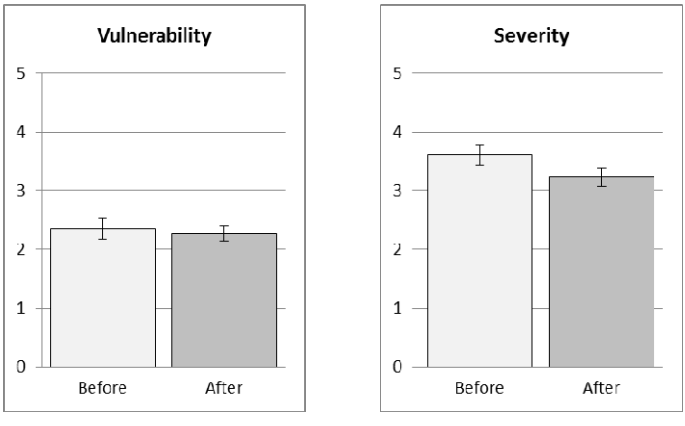


Fig. 4. Risk perception: vulnerability and severity, before and after the experience. Capped vertical bars denote ±1 SE.

7 Conclusions

To the best of our knowledge, our research is the first to study persuasive effects of simulated experiences of risk in the aviation safety education domain. Moreover, with respect to previous studies of simulated risk experiences, we have extended our investigation to important aspects such as self-efficacy and knowledge acquisition.

Overall, the experiment showed that serious games that simulate risk experiences can be a very effective tool for changing attitudes concerning personal safety topics, as well as for learning purposes: just playing a game level for 2-3 minutes resulted in a considerable improvement of users' self-efficacy and knowledge.

The fact that the virtual experience slightly decreased perception of risk severity seems to be consistent with the relation between self-efficacy and severity perception

that was pointed in the analysis, but also with the conclusions of our previous study [8] in which we recommended to explicitly depict human suffering and death in emotional ways if the purpose of the simulation is to increase risk perception. To clarify more thoroughly the effects of that recommendation, in future studies we will explore the possibility of making the aircraft evacuation simulation more threatening and increasing negative affect to explore if this would result in more persuasion or could instead be detrimental, resulting in defensive reactions of participants.

To further confirm the more than encouraging results we obtained, we are now planning an experiment with a larger sample of users in which we will contrast the effectiveness of the serious game vs. traditional safety cards and briefings. Although the studies reported in the literature are pessimistic about the effectiveness of safety cards and briefings (as discussed in Section 2), we believe a comparative analysis is needed, also to better quantify the advantages of the serious game solution. Finally, we will consider a longitudinal study to assess attitude and knowledge retention over time as a result of playing the game.

Acknowledgements. Michele Zanolin greatly helped in the development of the game and in setting up the evaluation.

References

1. Airbus: A320 Airplane Characteristics for Airport Planning, Technical Manual, http://www.airbus.com/support/maintenance-engineering/technical-data/aircraft-characteristics/
2. Bandura, A.: Social foundations of thought and action: A social cognitive theory. Prentice-Hall, Englewood Cliffs (1986)
3. Bandura, A.: Self-efficacy: The exercise of control. W.H. Freeman, New York (1997)
4. Bandura, A.: Social cognitive theory: An agentic perspective. Annual Review of Psychology 52, 1–26 (2001)
5. Boeing: Statistical summary of commercial jet airplane accidents. Boeing, Seattle, WA (2011), http://www.boeing.com/news/techissues/pdf/statsum.pdf
6. Chang, Y., Yang, H.: Cabin safety and emergency evacuation: Passenger experience of flight CI-120. Accident Analysis and Prevention 43, 1049–1055 (2011)
7. Chang, Y., Liao, M.: The effect of aviation safety education on passenger cabin safety awareness. Safety Science 47, 1337–1345 (2009)
8. Chittaro, L., Zangrando, N.: The Persuasive Power of Virtual Reality: Effects of Simulated Human Distress on Attitudes towards Fire Safety. In: Ploug, T., Hasle, P., Oinas-Kukkonen, H. (eds.) PERSUASIVE 2010. LNCS, vol. 6137, pp. 58–69. Springer, Heidelberg (2010)
9. Corbett, C.L., Cosper, D.K.: Effective Presentation Media for Passenger Safety I: Comprehension of Briefing Card Pictorials and Pictograms. Final Report DOT/FAA/AM-08/20, Federal Aviation Administration (2008)
10. Cosper, D., McLean, G.: Availability of passenger safety information for improved survival in aircraft accidents. Technical Report DOT/FAA/AM-04/19, Federal Aviation Administration, Washington, DC (2004)
11. de Hoog, N., Stroebe, W., de Wit, J.B.F.: The processing of fear-arousing communications: How biased processing leads to persuasion. Social Influence 3(2), 84–113 (2008)

12. Edwards, M.: Stress, behavior, training and safety. Flight Safety Foundation, Cabin Crew Safety 25(3), 1–5 (1990)
13. Fogg, B.J.: Persuasive Technology: Using Computers to Change What We Think and Do. Morgan Kaufmann, San Francisco (2003)
14. Fogg, B.J.: A Behavior Model for Persuasive Design. In: Proceedings of Persuasive 2009: 4th International Conference on Persuasive Technology. ACM Press (2009)
15. Grau, R.G., Martínez, I.M., Agut, S., Salanova, M.: Safety Attitudes and Their Relationship to Safety Training and Generalised Self-Efficacy. International Journal of Occupational Safety and Ergonomics 8(1), 23–35 (2002)
16. Katz-Navon, T., Naveh, E., Stern, Z.: Safety self-efficacy and safety performance. International Journal of Health Care Quality Assurance 20(7), 572–584 (2007)
17. Khaled, R., Barr, P., Aumann, Y., Fischer, R., Biddle, R.: Fine Tuning the Persuasion in Persuasive Games. In: de Kort, Y.A.W., IJsselsteijn, W.A., Midden, C., Eggen, B., Fogg, B.J. (eds.) PERSUASIVE 2007. LNCS, vol. 4744, pp. 36–47. Springer, Heidelberg (2007)
18. Leach, J.: Why people 'freeze' in an emergency: temporal and cognitive constraints on survival Responses. Aviation, Space and Environmental Medicine 75(6), 539–542 (2004)
19. Leach, J.: Cognitive paralysis in an emergency: the role of the supervisory attentional system. Aviation, Space, and Environmental Medicine 76(2), 134–136 (2005)
20. Luszczynska, A., Schwarzer, R.: Multidimensional health locus of control: Comments on the construct and its measurement. Journal of Health Psychology 10, 633–642 (2005)
21. Meijnders, A., Midden, C., McCalley, T.: The Persuasive Power of Mediated Risk Experiences. In: IJsselsteijn, W.A., de Kort, Y.A.W., Midden, C., Eggen, B., van den Hoven, E. (eds.) PERSUASIVE 2006. LNCS, vol. 3962, pp. 50–54. Springer, Heidelberg (2006)
22. Muir, H., Thomas, L.: Passenger education: past and future. In: Proceedings of the 4th Triennial International Aircraft Fire and Cabin Safety Research Conference (2004)
23. NTSB (National Transportation Safety Board): Emergency Evacuation of Commercial Airplanes. Safety Study NTSB/SS-00/01. NTSB, Washington, DC (2000)
24. Proulx, G.: Playing with fire: Why People Linger in Burning Buildings. ASHRAE Journal 45(7), 33–35 (2003)
25. Rogers, R.W.: Cognitive and physiological processes in fear appeals and attitude change: A revised theory of Protection Motivation. In: Cacioppo, J.T., Petty, R.E. (eds.) Social Psychophysiology: A Sourcebook, pp. 153–176. Guilford Press, New York (1983)
26. Thomas, L.J.: Passenger Attention to Safety Information. In: Bor, R. (ed.) Passenger Behaviour, Ashgate, UK (2003)
27. Schwarzer, R., Jerusalem, M.: Generalized Self-Efficacy scale. In: Weinman, J., Wright, S., Johnston, M. (eds.) Measures in Health Psychology: A User's Portfolio, Causal and control beliefs Windsor, pp. 35–37. Nfer-Nelson, UK (1995)
28. Slovic, P., Peters, E.: Risk Perception and Affect. Current Directions in Psychological Science 15, 322–325 (2006)
29. Van Gerwen, L.J., Spinhoven, P., Diekstra, R.F.W., Van Dyck, R.: People who seek help for fear of flying: typology of flying phobics. Behavior Therapy 28(2), 237–251 (1997)
30. Zaalberg, R., Midden, C.: Enhancing Human Responses to Climate Change Risks through Simulated Flooding Experiences. In: Ploug, T., Hasle, P., Oinas-Kukkonen, H. (eds.) PERSUASIVE 2010. LNCS, vol. 6137, pp. 205–210. Springer, Heidelberg (2010)
31. Zyda, M.: From visual simulation to virtual reality to games. IEEE Computer 38(9), 25–32 (2005)

Towards Persuasive Technology for Software Development Environments: An Empirical Study

Ingo Pribik[1,2] and Alexander Felfernig[1]

[1] Institute for Software Technology, Applied Software Engineering,
Graz University of Technology, Austria
[2] Flextronics, Austria
ingo.pribik@at.flextronics.com,
alexander.felfernig@ist.tugraz.at

Abstract. This paper describes an empirical study of a developed persuasive software tool (PerSoDeMetrics) and its effects on the software engineering area. The duration of the study was ten months. The persuasive tool was introduced for a small development team in the real-world. The study tries to analyze the effect of this tool being applied by software developers. The major goal was to persuade software developers to focus on the improvement of their software components. Based on this challenge we developed a plug-in which includes persuasive technology mechanisms. The persuasive software tool is a plug-in/extension in the development environment for Visual Studio 2010. It calculates the software metrics and provides recommendations as how developers can improve the quality of their software components.

Keywords: Persuasive Technology, Software Quality, Software Usability, PSD model.

1 Introduction

Software development teams are under high pressure to release software products within limited time frames with all the required features and high software quality. Usually the software quality suffers due to time pressure. However, quality has an important role in software engineering. Maintainability and extendibility of software products require a high software quality. These are the basic expectations for software libraries to be able to use them for different software projects. But how can we describe, identify, analyze a software quality and at the same time try to motivate developers to improve their software artefacts? Software quality is often expressed as follows: *I know it when I see it. Terms such as good quality, bad quality* [4]. We know that there are more dimensions which influence software quality. Brain Henderson [3] defines software quality as follows: *reliability*, *availability*, *maintainability* and *usability*. We focused on improving the maintainability of software products. The other aspects like *reliability*, *availability* and *usability* are not included in our study. Maintainability depends on *understandability*, *modifiability*, and *testability* [3]. Software metrics [6, 7, 8, 20] are one of the most common used

M. Bang and E.L. Ragnemalm (Eds.): PERSUASIVE 2012, LNCS 7284, pp. 227–238, 2012.

technologies to identify and analyze the software complexity and software design. On the market, there are applications available which calculate and analyze software metrics [5]. As an example, *SourceForge* is a plug-in for Eclipse which calculates software metrics. Microsoft introduced the *Code Metrics* for their development environment Visual Studio. But even when such a tool is available in the software development environment, developers often still ignore the quality of their software components. The missing link is to use persuasive technologies for such a kind of tool. This technology could support the improvement of the software quality by confronting programmers with complexity evaluation results. During the last years several persuasive applications were being developed with the scope to change the behavior and attitudes of the user. In this context, persuasive technology is defined as *technology that is designed to change attitudes of behaviors of the users through persuasion and social influence, but not through coercion* [1].

The major application domains of persuasive technologies are health [16], fitness [12], e-commerce or games [11]. This paper presents an implementation of a *behavior change support system* (BCSS) [23] for software development environments (in our case Visual Studio 2010).

2 Motivation and Evaluated Environment

For the purpose of our evaluation we have chosen a development team consisting of six developers located in Malaysia. The Malaysian team has already started to work on a software package (half of the scheduled project time already passed). The software product which has been developed by the Malaysian team is the next generation of a shop floor control system. This software product/library is used to cover the traceability requirements. The main requirement of traceability is to collect all predefined data down to the lowest components on each process step necessary and to support an optimal operational flow for the products. These data allow tracing back what kinds of different components were assembled in semi-finished or finished products. Different development teams, for example the customization teams are going to use this traceability library. They have added features for several projects in order to be able to support optimal process/operation flow in the production. Optimal operational efficiency is a goal for an electronics manufacturing service (EMS). Therefore, in terms of extendibility and maintainability, the software quality is an important factor of successfully implemented projects. From the software reuse [18] perspective, important customized software components will be included in the next traceability library release. Development activities such as customization and software reuse have to be completed quite often in a short time period. In such a case the developers do not see the need to improve the quality of the methods. They are adding the methods and classes into the library with the focus on a high output. The plug-in for Visual Studio 2010 with persuasive technology should try to change the behavior of the developers to be more sensitive to their software quality. The tool is also foreseen to help the developers during the refactoring process [17]. The highlight of methods and classes which do not fulfill predefined quality criteria are the basis of software quality improvement. Knowing what kind of methods and classes should be improved we have defined and introduced a software metric suite (Figure 2). In our

empirical evaluation our main focus was to encourage developers to take care of the quality of committed methods. The underlying assumption is that a redesign of classes would cost too much and would be too risky for the whole software package.

2.1 Defined Software Metric Suite

The software metrics can be classified based on Fenton and Pfleeger [2] in *process* metrics, *product* metrics and *resource* metrics. For our empirical study we focused on the *product* metrics with the following metric suite (Figure 1).

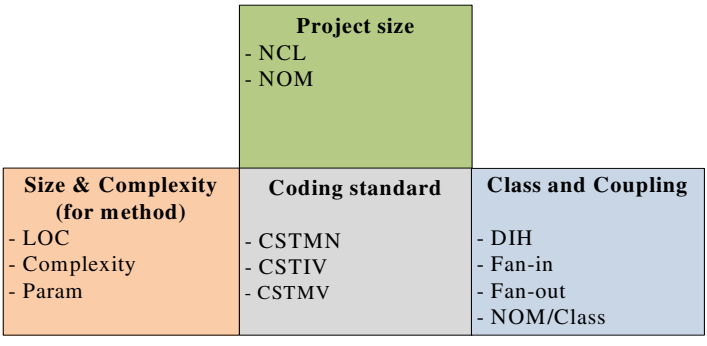


Fig. 1. The metric suite consists of four metric areas project size, size and complexity, coding standard and class/coupling. The project size metrics informs about the system growth during per iteration. The size and complexity metric evaluates the methods. The coding standard metric checks if the predefined format for the coding style was followed. The coupling metric is an input to analyze the class design.

The metric suite is a collection of different metric suites [3, 5]. We added an individual metric collection "Coding standard" to.

The following metrics have been integrated in our persuasive environment.

Project size:
NCL: The metric counts the number of classes of the system.
NOM: It counts the number of methods implemented in the system.

Size & Complexity for methods:
Complexity: The *complexity* will be calculated based on McCabe [13]. This metric counts the different code paths of a method.
LOC: The *Lines of Code* metric counts the lines of code per method. Blank lines and comment lines will be ignored.
Param: The *param* metric counts the number of parameters per method. Return parameters will be ignored.

Coding standards:
CSTMN: The *code standard for method name* checks if all methods start with an upper case letter.

CSTIV: The *code standard for instance variable* checks whether all instance variables start with an underline and lower case.
CSTMV: The *code standard for method variable* checks whether all methods variables start with lower case.

Class & Coupling:
DIH: The *depth of inheritance* calculates the inheritance of user classes and ignores system classes.
Fan-in: The Fan-in [3] counts how many classes are instantiated in a class. The instantiation of a class will be counted only one time.
Fan-out: The Fan-out [3] metric calculates how many classes are going to be instantiated by a given class.
NOM/Class: The metric *number of methods* counts the methods per class.

2.2 Used Threshold for the Metrics

We determined the measuring range for each metric into three areas. These areas are named *good area*, *critical area* and *very critical area*. We defined methods or classes which are marked as critical are able to maintain. The developers will still be informed by our tool to improve them. Methods or classes which are marked as very critical have to be improved because the maintenance costs of those software artifacts are too high. The predefined thresholds are identifying to which area a class or method belongs to. The list of used metrics and their thresholds are listed in (Table 1).

Table 1. Used thresholds for software metric types

Metric type	Good area	Critical area	Very critical area
LOC*	<= 29	Between(29, 74)	>= 74
COMPLEXITY*	<= 5	Between(5, 14)	>= 14
PARAM*	<= 4	Between(4, 10)	>= 10
DIT*	<= 3	Between(2, 8)	>= 8
FANIN*	<= 9	Between(9, 56)	>= 56
FANOUT**	<= 2	Between(2, 5)	>= 5
NOM/Class**	<= 18	Between(18, 26)	>= 26

* These thresholds are the outcome of an evaluation of 100 object-oriented systems written in C# and Java [9].

** An internal software project was used to calculate these thresholds. For the critical threshold we used the (avg + 2stdv) and for the very critical we used the (avg + (3stdv)).

3 Implemented Persuasive Environment

Different process models [14, 19] are available to design a persuasive system. The persuasive system design PSD model [19, 10] is a framework that we used as a guide to design our persuasive tool. The PSD model analyzes in detail the intent, event and

strategy of persuasion [24]. The *persuader* who supported this tool is commercial company. First of all we defined the overall behavior change we wanted to achieve with our BCSS tool. The first target of the behavior change is that the developer should check the quality of his/her software artifacts and improve them at least once a week. This behavior change is also known as *macrosuasion* goal [1]. With our approach the complexity of the methods should be decreased resulting in an improved maintainability. The second target was that the developers follow our coding standard. The readability of software code depends on coding standard and is a maintainability dimension. There is no doubt that the acceptance of such a tool depends on usability. Additional the PSD model describes software characteristics for the BCSS's under for categories, *primary task support*, *computer-human dialogue support*, *perceived system credibility* and *social influence* [24]. Therefore, we made some usability tests to come out with a tool which should be intuitive, self-explained and easy to use. B.J. Fogg defines three dimensions relevant to achieve a behavior change. These three dimensions are *trigger*, *availability* and *motivation* [5]. The behavior change support system (BCSS) should give the developer the availability to improve his/her software components ad hoc and this should motivate them to change his/her behavior.

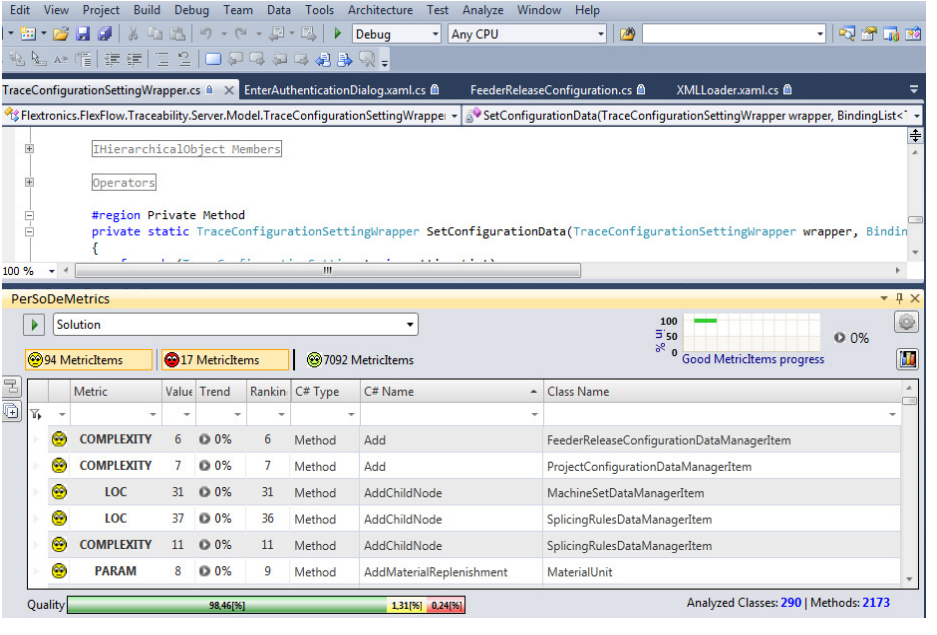


Fig. 2. The persuasive metric tool in the Visual Studio 2010 environment

The persuasive tool/BCSS is developed in C# and is a view in the development environment for Visual Studio 2010 shown in (Figure 2). The design principles for our tool were reduction, tunneling, self-monitoring and suggestion as our persuasive elements [1]. The tool also has a lot of microsuasion [1] elements. Depending on the metric thresholds measured metric values can belong to the *good*, *critical* or *very*

critical area. This information will be displayed with different emotion icons (Figure 4). The metric ranking information for each metric should give the information of how good or how bad the calculated metric value is related to the other metric results of the same metric type (Figure 4). A column trend visualizes the progress of each development change if it was good or bad (Figure 4).

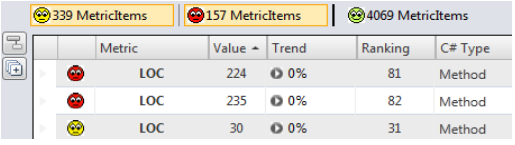


Fig. 3. This is the detail view of the persuasive tool. Above the grid the numbers of calculated metrics are being displayed and classified into the three areas (*critical*, *very critical* and *good*). The column in the grid includes detail information about each calculated metric type, calculated value, the current trend, ranking and if the metric is a method metric or class metric.

All this displayed information will be refreshed after each metric calculation. The displayed data can be seen as self-monitoring information. It shows the actual quality status and should motivate the developer to improve it. With a click on one of the metric rows in the grid (Figure 2) the persuasive tool opens a recommendation window shown in (Figure 4).

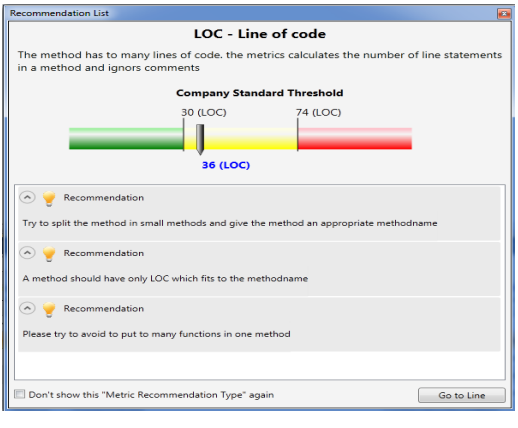


Fig. 4. Guide window with self-monitoring, suggestion and tunneling concept

This window shown in (Figure 4) includes the self-monitoring information of the selected metric again more in detail. It includes also a short description how the metric will be calculated. The button “Go to Line” directly guides the developer to the marked software artifact which can be seen as a reduction [1] persuasive element. As an additional strategy to the BCSS we send a metric report to the developers first per iteration and then weekly based. The strategy of the PSD model consists of two elements, *message* and *route* [24].

4 Results of Empirical Study

As already mentioned the BCSS tool was introduced for a small development team in Malaysia consisting of six developers. The development team uses the iterative development approach. The schedules for iterations have been defined within six weeks. A set of new functions was assigned to the developers to implement per iteration. Our *behavior change support system* was introduced after the iteration 6 (I06).

Iteration 01 – Iteration 06:
In these iterations the developers did not use the software metric tool.

Iteration 07 – Iteration 09:
In these iterations the behavior change support tool has been applied to the developers.

Iteration 10 – Iteration 11:
In these iterations the developers did not apply or at least partially applied the software metric tool. They had struggled hard to implement missing features in this short time frame.

Iteration 12- Iteration 14:
In these iterations developers again used the software metric tool to improve the software quality.

4.1 Software Metric Trend

With the trend of the software metrics we want to visualize how the behavior of the developers changed with and without using the BCSS tool.

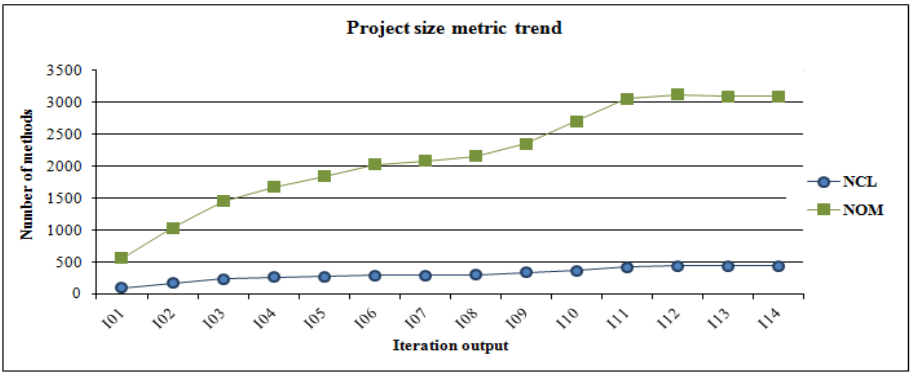


Fig. 5. The graphic shows the trend of the project size. It includes the *number of methods* and *number of classes* for the whole traceability module.

All metric graphics include the GUI (Graphic User Interface) and the BLL (Business Logic Layer) of the project. According to the graph, we can see a constant

growth of the project size until iteration 12. All known features have been implemented by the developers. From iteration 12 to 14 developers were focusing on refactoring and improving software quality. The method metric trend (Figure 6) shows an improvement after the use of the BCSS tool in iteration 07. From iteration 07 to iteration 09 the developer added the approximate same amount of methods based on the previous iterations.

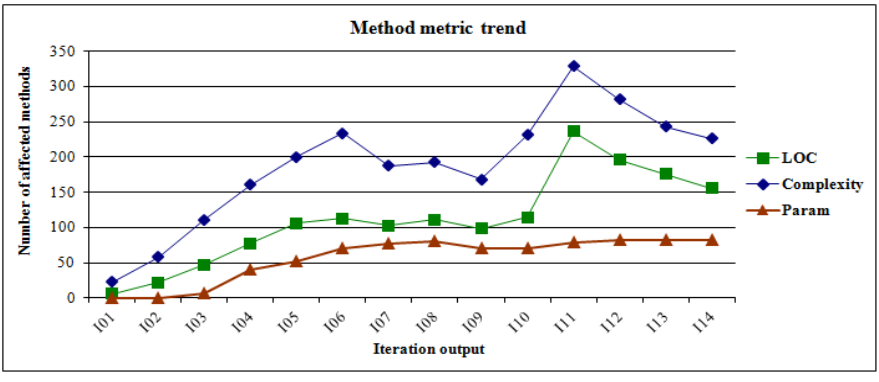


Fig. 6. The graphic shows all affected methods which have been identified by at least one method metric. An affected method will only be counted once even it includes several *critical or very critical* method metrics. Method metric are (LOC) *Lines of Code* metric, (Complexity) metric measures the different code paths, or (Param) metric which counts the number of input parameters per method.

That means the developers improved the software quality of the existing methods too. Between iteration 10 and 11 the users/developers rarely used the metric tool because they were under high pressure to add a lot of features in a short time period.

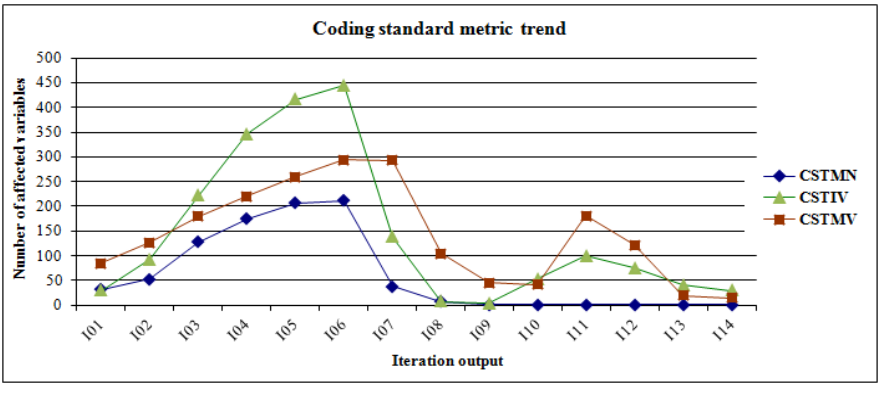


Fig. 7. The graphic shows all affected methods which have been identified by the coding standard metric. CSTMN stands for *coding standard for method name*, CSTIV stands for *coding standard instance variable* and CSTMV stands for *coding standard for method variable.*

In iteration 12 we sent out a software quality status per e-mail to each developer on a weekly basis. This was used as an additional trigger for the developers to motivate them. We also informed the developers, that they should focus on the LOC metric and coding standard metric in the next iterations to keep it simpler for them. Based on the output we can see that the developers could improve the quality but we should also take into account that after iteration 12 (Figure 5) no new features were assigned to the traceability module.

The best results we have in the area of the coding standard metrics (Figure 7) after the release of our BCSS tool. The trend shows that the developers took better care of the software metric and mostly solved all of the highlighted issues. We can assume the main reason of this trend lies in the simplicity to improve the metric values.

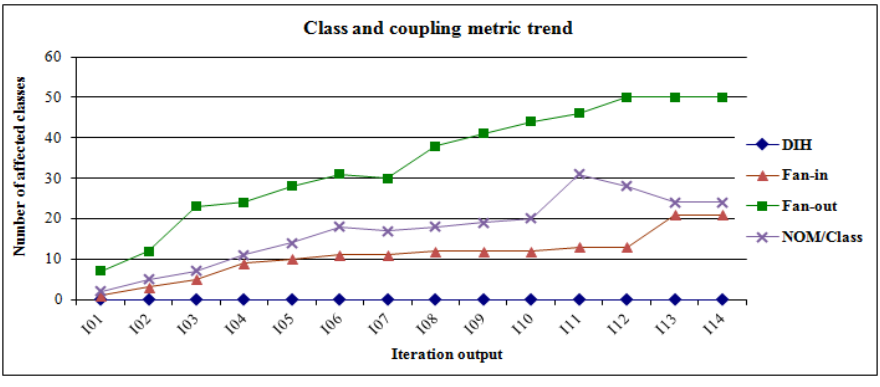


Fig. 8. The graphic shows all affected classes based on the class and coupling metrics. A class will only be counted once when one *critical* or *very critical* metric was assigned to it. The used class and coupling metrics are *depth of inheritance* metric, *Fan-in* metric, *Fan-out* metric and *number of method per class* metric.

The class and coupling metric were not improved by the developers. The reason for that would be to make a redesign of the system which was to risky for them at this time. We can see an improvement on the NOM/Class metric which counts the number of methods per class. The Fan-out metric has a constant growth through all iterations. That means we have some classes which are instantiated very often in other classes. The coupling has big influence in the complexity of the system. Developers have to understand all coupled classes of a class when they would like to change or extend the class. Therefore we should try to keep the Fan-out value as low as possible [21, 22]. The software metric tool tried at least to identify and inform the developers of the coupling situations.

4.2 Use of the Software Metric Tool

There are several parameters responsible to get this kind of software quality improvement which is shown above. To be sure that the BCSS tool was one of them we tried to get the information of how often the tool was used by the developers. We

implemented trace information in the tool and distributed it to each developer again. We monitored the click events of the developers of the BCSS tool. The first week per iteration was used to analyze the new assigned requirements and to define the tasks for each developer. The table (Table 02) shows the last five weeks of the I10 (iteration 10). Although the developers were under pressure during the I10, they tried to use the BCSS tool.

Table 2. Software metric tool use per user/developer

User	week01	Week02	Week03	Week04
User01	6	0	171	5
User02	12	11	0	2
User03	1	64	0	5
User04	0	31	0	18
User05	3	16	45	1

They introduced from I09 to I10 based on the (Figure 5) 351 new methods and 34 new classes. In the I10 they only could keep the performance of the coding standard metric shown in (Figure 7). Based on this result we can assume that the tool has contributed to the software quality improvement.

4.3 User Performance

With the user performance we are going to show how each user/developer improved his committed software components per iteration. We could only retrieve the history information from I04 (iteration 04) onwards from our team foundation system TFS system.

Table 3. Committed software quality performance per user/developer for each iteration

User	I04	I05	I06	I07	I08	I09
User01	78,83%	81,88%	79,27%	78,80%	80,78%	90,67%
User02	83,44%	88,93%	90,87%	95,20%	93,92%	94,01%
User03	80,28%	45,33%	81,19%	81,30%	82,36%	81,41%
User04	64,10%	89,02%	88,33%	94,82%	97,53%	96,92%
User05	88,07%	49,58%	74,25%	74,12%	68,18%	85,38%
User06	77,33%	83,56%	81,09%	87,54%	71,43%	-

The performance calculation (Table 3) per user was calculated based on the performance calculation (1). The calculation considers changed or new developed methods per iteration. The variable x is the number of all changed or new developed methods per user. The variable y is the number of all methods per user which have at least one critical or very critical metric.

$$(100 / x) * y = \text{performance} . \quad (1)$$

We can interpret, that there is an improvement from iteration 06 to iteration 09 of each user/developer. The user01/developer01 recorded the best improvement from

iteration 06 to iteration 09 which is 11,4%. Only the user06/developer06 could not improve his performance based on the committed history information. The average improvement from iteration 06 to 09 for user01 to 05 is 6,896%.

5 Conclusions and Future Work

One of our main targets was to show how a small software development team can be influenced by a persuasive technology tool. The PSD model helped us about the design principles for the BCCS. We could show that the software quality could be improved as a direct consequence of applying the presented persuasive technologies. Of course the trigger should be improved to get a higher intension by each developer. The trend of the monitored software metrics and the user performance shows us that the behaviors of the developers were changed. The next stage will be to use this tool in the context of other development projects. Another interesting question would be to study how such behavior change support tool will influence and motivate developers in different locations/cultures [15]. We have already started to implement this tool for a small development team in Europe. A second idea for improving the motivation of developers is to design a webpage accessible for each developer, showing, the current software metrics of a project where they have been assigned to.

References

1. Fogg, B.J.: Persuasive Technology. Morgan Kaufmann, Boston (2003)
2. Fenton, N., Pfleeger, S.: Software Metrics: A Rigorous and Practical Approach, 2nd edn. International Thomson Computer Press, London (1996)
3. Henderson-Sellers, B.: Object-Oriented Metrics: Measure of Complexity. Prentice Hall, Upper Saddle River (1996)
4. Kan, S.H.: Metrics and Models in Software Quality Engineering, 2nd edn. Addison-Wesley Publishing Company (2008)
5. Lanza, M., Marinecsu, R.: Object-Oriented Metrics in Practice. Springer (2006)
6. Jamali, S.M.: Object Oriented Metrics. A survey approach Technical report, Department of Computer Engineering, Sharif University of Technology, Tehran, Iran (2006)
7. Schroeder, M.: A Practical Guide to Object-Oriented Metrics. IT Professional – IEEE 1(6), 30–36 (1999)
8. Martin, R.: OO Design Quality metrics – An Analysis of Dependencies. Position paper, Proc. Workshop Pragmatic and Theoretical Directions in Object-Oriented Software Metrics, OOPSLA 1994 (October 1994)
9. Alves, T.L., Ypma, C., Visser, J.: Deriving metric threshold from benchmark data. In: IEEE International Conference on Software Maintenance, ICSM 2010, pp. 1–10 (2010)
10. Torning, K., Oinas-Kukkonen, H.: Persuasive system design: State of the art and future directions. In: Proceedings of 5th International Conference on Persuasive Technology, Claremont, CA, USA. ACM International Conference Proceeding Series, vol. 350, pp. 26–29 (2009)
11. Gram-Hansen, L.B.: Geocaching in a Persuasive Perspective. In: Proceedings of the 4th International Conference on Persuasive Technology, article 34 (2009)

12. Harjumaa, M., Segerstahl, K., Oina-Kukkonen, H.: Understanding Persuasive software functionality in practice: a field trail of polar FT60. In: Proc. 4th Internal Conference on Persuasive Technology, Claremont CA, US. ACM International Conference Proceeding Series, vol. 350, pp. 1–9 (2009)
13. McCabe, T.: A complexity measure. IEEE Transactions on. Software Engineering SE-2, 308–320 (1976)
14. Fogg, B.J.: Creating Persuasive Technologies: An Eight-Step Design Process. In: Proceedings of the 4th International Conference on Persuasive Technology, article 44 (2009)
15. Kahled, R., Biddle, R., Noble, J., Barr, P., Fischer, R.: Persuasive interaction for Collectivist cultures. In: Proceedings of the 7th Australasian User Interface Conference, pp.73–80 (2006)
16. Chi, P.-Y(P.), Chen, J.-H., Chu, H.-H., Lo, J.-L.: Enabling Calorie-Aware Cooking in a Smart Kitchen. In: Oinas-Kukkonen, H., Hasle, P., Harjumaa, M., Segerståhl, K., Øhrstrøm, P. (eds.) PERSUASIVE 2008. LNCS, vol. 5033, pp. 116–127. Springer, Heidelberg (2008)
17. Fowler, M.: Refactoring: Improving the Design of Existing Code. Addison-Wesley, Boston (1999)
18. Krueger, C.W.: Software reuse. ACM Computing Surveys 24, 131–184 (1992)
19. Oinas-Kukkonen, H., Harjumaa, M.: Persuasive Systems design: Key Issues, Process Model, and System Features. The Communications of the Association for Information Systems 24, article 28, pp. 485–500 (2009)
20. Braind, L., Daly, J., Porter, V.J., Page, J.: A Comprehensive Empirical Validation of Product Measures in Object-Oriented Systems. In: Proc. 5th International Symposium Software Metrics, Bethesda, Md, pp. 246–257 (November 1998)
21. Page, J.: The Practical Guide to Structured Systems Design, 2nd edn. Prentice Hall International, Englewood Clifss (1988)
22. Card, D.N., Glass, R.L.: Measuring Software Design Quality. Prentice Hall, Englewood Clifss (1990)
23. Oinas-Kukkonen, H.: Behavior Change Support Systems: The Next Frontier for Web Science. In: Proceedings of the Second International Web Science Conference (WebSci 2010), Raleigh, NC, US, April 26-27 (2010)
24. Oinas-Kukkonen, H.: Requirements for Measuring the Success of Persuasive Technology Applications. In: Proceedings of the 7th International Conference, New York (2010)

Persuasive Technology Considered Harmful? An Exploration of Design Concerns through the TV Companion

Wolfgang Reitberger, Florian Güldenpfennig, and Geraldine Fitzpatrick

Institute for Design & Assessment of Technology, Vienna University of Technology, Argentinierstraße 8, 1040 Vienna, Austria
{wolfgang.reitberger,florian.gueldenpfennig, geraldine.fitzpatrick}@tuwien.ac.at

Abstract. Persuasive Technology has great potential to positively influence users in a wide variety of areas such as health, well-being and environmental sustainability. However, concerns have been raised regarding issues such as surveillance, lack of autonomy and coercion that might be involved with the application of such technologies. This paper articulates a set of design concerns around these themes and presents a study of the TV Companion, an instance of a behavior change system that takes into account these concerns. The TV Companion is an application aimed to address the societal issue of prolonged TV consumption with its related potential negative effects and aims to foster more reflective TV viewing among children and their parents. To explore the application of these design approaches in a real world setting we report the findings of a field study of the TV companion in three families with pre-school children and reflect on how the design concerns shaped the experience and use of the TV companion.

Keywords: design concerns, autonomy, surveillance, coercion, TV viewing, families, prototype, field study.

1 Introduction

Over the last years, Persuasive Technology (PT) has successfully been applied to positively influence users in a wide variety of areas such as health [1], well-being [2, 3] and environmental sustainability [4]. Nevertheless, these technologies target phenomena at the core of what we are and how we see ourselves and the world, namely our attitudes and behaviors. Thus, not surprisingly concerns over PT and its applications have been raised over the last years. These concerns include the preservation of autonomy and free choice, the dichotomy between coercion and reflection and the risk the utilization of surveillance poses for the users' privacy.

The TV Companion is an application that we designed taking into consideration the design concerns for behavior change systems, outlined above, in a concrete application. It can be seen as an instance of a new class of behavior change systems which is rooted in the tradition of Persuasive Technology but also takes into account some of the criticisms and concerns that have been brought forward. The TV Companion

M. Bang and E.L. Ragnemalm (Eds.): PERSUASIVE 2012, LNCS 7284, pp. 239–250, 2012.

addresses the topic of TV consumption among children and aims to nudge parents and their children towards a critical reflection of their TV watching habits.

Excessive TV consumption constitutes a societal problem with a variety of related negative health and cognitive effects. A recent Australian study shows that the amount of TV watched there in 2008 reduced life expectancy by 1,8 years for men and by 1,5 years for women. Even more dramatically, those who watched an average of 6 hours per day had a reduction of their life expectancy by 4,8 years, making TV a risk factor comparable with lack of physical activity and obesity [5]. Furthermore, there is evidence that television, and particularly fast paced cartoons, has negative effects on pre-school children's working memory and self regulation [6]. Based on a review of the possible harmful consequences television viewing has for children, the American Academy of Pediatrics recommends to limit TV exposure to between one and two hours of quality programming per day for everyone above the age of two years [7].

We thus designed the TV Companion application to specifically influence the TV watching behavior of younger children in order to provide them and their parents with a tool to help them avoid the negative effects of prolonged television consumption. We then study its use in field deployments with three families.

This paper starts with a presentation of our three design concerns, pointing out issues being raised in literature and an alternative design response, and continues with a description of the prototype including the design rationale that addresses the issues in a different way. After this the study setup and results are outlined. This is followed by a discussion, some concluding remarks and an outlook towards future work. The main contribution of this work is twofold: the identification of three design concerns affecting PT systems and their implementation; and reflections on how the principles of these concerns can be realized in a prototype and experienced in use.

2 Design Concerns

Based on an analysis of the literature in PT and related areas we identified three main design concerns for behavior change systems. We propose that in order to avoid the possible harmful consequences of such systems, these concerns should be considered when designing Persuasive Technology and suggest alternative design responses.

2.1 Autonomy and Free Choice

Persuasive technology has been criticized for potentially infringing on the autonomy of its users [8]. More specifically, DiSalvo et al. [9] point out that in PT applications, the desired behavior is determined by the designer, thus often leaving little space to accommodate the values and ideas of the users. The importance of taking into account human values in the design process has been emphasized by Friedman et al. [10]. One could argue that the persuasive strategy of tailoring, which has been proposed by Fogg [11], and more recently has also been included in the Persuasive Systems Design (PSD) model [12], would encompass adapting the means and ends of

such systems to the individual users. Yet, both describe this strategy only to include the adaption of the means employed to persuade users and not of the ends of persuasive interventions. This means-based tailoring or adaptation approach is also taken in the area of persuasion profiling, where different persuasive strategies are applied based on the preferences of specific users [13]. As Fogg [11] points out, tailoring increases the persuasive potential of system, which is supported by the findings of Kaptein et al.[13].

In order to alleviate the aforementioned concerns regarding autonomy and to foster the self determination of the users, we propose to extend this strategy to also include ends-based tailoring. In the design of systems based on this strategy, not only the presented information would be tailored to "the potential needs, interests, personality, usage context, or other factors relevant to a user group."[12], but also the end, the desired behavior change. This end should not be completely determined by the designer. Instead, it should also be able to reflect the goals, values and ideas of the individual user.

2.2 Coercion versus Reflection

In his book defining the area of PT Fogg already points out that there is a fine line between persuasion and what could be considered coercion [11]. He emphasizes that coercion involves force, whereas persuasion denotes voluntary change in attitude or behavior. Despite this clarification, the field and the concrete PT applications that were developed over the last years have been criticized for leaning too much towards coercion, ultimately leading to designs fostering technological control over the user [14].

We thus suggest that the designers of PT should even more actively try not to cross the line towards coercion and build enough "wiggle room" for the users into theirs systems. Concurring with the argument, that the final and ideally rational decision whether to adapt a new behavior or not should be left to the user [8], we argue that PT systems should not rely on force but rather promote reflection of the users' own actions in order to help them to reach the desired behavior. One strategy to reach this aim is to utilize the perCues framework [15], which implies the creation of awareness for a certain behavior by increasing its visibility, allowing the user to reflect about it and change the behavior if she/he desires to do so.

2.3 Surveillance and Privacy

Closely related to the first two design concerns emphasizing autonomy and the reflection of one's own behavior is the concern for the users' privacy. One of the (acceptable) principles of persuasion according to Fogg [11] is using surveillance. The only caveat he points out is that the surveillance has to be overt, even though there is the concern that surveillance may lead to "public compliance without private acceptance" [11] leading only to superficial and unsustainable behavior change. Reverting to the notion of the panopticon, Jespersen et al. [16] highlight the dangers of persuasion with the means of technologies which can observe and track users and their behavior

at any given time, leading to constant and mass surveillance. They point out that there is an ever growing amount of surveillance in our society, leading to what one could term a surveillance society. This sentiment is echoed by [14], who depict a rather dystopian view of a "persuasive society" which is governed by rational model of control and efficiency, leading to discipline through surveillance and the looming threat of punishment as discussed by Foucault [17].

For the design of Persuasive Technologies this implies that merely making the surveillance known to the users is not enough to answer this concern, since, as shown in the Panopticon, this could even lead to additional negative effects. Designers should not only make surveillance overt, but also carefully consider the amount of it used by their application. Furthermore the users should be empowered to actively control an individual level of privacy restriction imposed by an application. For instance, not every aspect of a user's behavior needs to be broadcasted immediately and inevitably via social media channels.

3 TV Companion Prototype

In this chapter we describe the design rationale and the technical setup of our prototype – the so-called *TV Companion* (see figure 1). In the preceding section we introduced three design concerns, namely *autonomy and free choice*, *coercion versus reflection* and *surveillance and privacy*. We explain how we carefully paid attention to address all three concerns when building the prototype. Before that, we present a short scenario to illustrate how the TV Companion could be used.

3.1 Using the TV Companion

In a typical usage scenario the user – here a child - decides how long she/he wants to spend watching TV and picks a number of RFID chips according to this amount of time (see figure 1 (b) and (c)). The RFID chips or tokens correspond to different amounts of time (15, 30, 45 minutes). After switching on the TV Companion device, the user places the tokens upon the box (figure 1 (a)). The user is then thanked by a voice for each token and the amount of time is read out aloud. The total amount of TV-minutes is cumulated from the single tokens. Time is counted back from the total time-amount to zero and counting starts immediately after a token has been read by the TV-Companion. When watching TV the user can estimate the remaining time by the color-coding of the LED on top of the device. Additional tokens, i.e. TV-minutes, can be brought in at any time. For instance, if the device is going to shut down soon, but the TV-broadcast hasn't finished yet, the user can throw in another token with 15 minutes. Finally, when all time is up, the TV-Companion starts blinking red, and the built-in infrared remote control turns off the TV.

In short the TV Companion is a device that helps the users keep track of their TV consumption in a playful way. However, as we will see when we later report the findings from our field studies with families who used the prototype, this does not happen in a rigorous and imposed fashion.

Fig. 1. (a) TV-Companion device (b+c) with RFID tokens. The device's electronics components comprise an Arduino microprocessor board, a RFID reader module, infrared TV remote control, a module for playing back wave recordings and speakers.

3.2 Design Rationale

We go on here to report how we designed the TV Companion to take specific account of each of our design concerns. Firstly, the TV Companion device is designed to grant multiple facets of *autonomy* to the users. One the one hand, the decision on how many RFID tokens (i.e. "TV minutes") the children should own is up to the parents. Also, the parents can decide on an appropriate time interval for RFID chip dispersion, e.g. on a daily or weekly basis, or per single TV show. Furthermore, the child is equipped with autonomy and freedom of choice as well. The children can apportion their TV minutes and save them for their favorite TV broadcasts. It is their decision to place a RFID chip on the wooden box or not. The box is also delivered as a plain wooden box and the children were encouraged to customize the box, aiming to give the children a sense of connection and ownership. In in-home workshops before the start of our field studies the children painted the wooden box of the TV Companion. In addition, they drew small pictures on their RFID tokens or chards (see figure 2 and 3).

Secondly, the act of deliberately putting the token on the TV companion leads to the design concern *coercion versus reflection*. Several features are built into the prototype, which aim at a deeper engagement of the user with his/her actions. In principle we hope that our device can support the children to reflect on their behavior by strengthening the appreciation for TV consumption as being about time, and about spending time and limited resources. In addition, we deliberately designed the device not to be "bullet-proof". This means that there is no built-in mechanism for a brute-force shutdown of the TV. The ideology of such a brute force and "merciless" control mechanism would stand in strong conflict with our design motivations and our belief that there can be a persuasive technology that grants the user *autonomy* and some

"wiggle room". With the TV Companion the device can be tricked and bypassed (e.g. by covering up the infrared LEDs to prevent the signal from transmission).

Fig. 2. Children painting the TV Companion

Finally, we chose not to include any explicit logging or *surveillance* features into our device. If the parents wanted to keep track of the TV consumption of their children they can accomplish this by means of counting the already used RFID tokens. This is part of the parents' autonomy and we as designers chose not to interfere with any aspects concerning the parents' or the child's privacy.

In summary, we hoped that the design of the TV companion left both the parents and the children with enough freedom and autonomy to determine both means and ends in keeping track of a healthy TV behavior.

3.3 Technical Description

The TV-Companion is realized using electronic prototyping components and off-the-shelf products (see figure 1). The centerpiece of the device is an *Arduino* microprocessor board, which controls an attached audio-board, a RFID reader module and an infrared remote control for turning off the TV (so-called *TV-B-GONE* remote control). Switching on the TV-Companion device enables the user to transfer "TV-minutes" to her/his time-account by means of RFID tokens associated with different amounts of minutes (15, 30, 45 min). When the account is used up, the infrared remote control will constantly try to shut down the TV.

As soon as a RFID token is put on the TV-Companion it is validated by the microprocessor and the amount of minutes associated with the token is added to the time account. A token cannot be used twice, i.e. it gets rejected after consumption, but it can be "recharged". Recharging works by placing a special "recharge token" on the device, which flags all tokens as unused in the internal database. For feedback, each action by the user is commented by an assigned audio recording. E.g., the TV-Companion is capable of announcing the time. The voice prompts are recorded by a friendly female speaker and accompanied by cheerful voices of children. In addition, the device features a multi-color LED lamp for giving visual feedback on the remaining time (the lamp turns from green to red as time passes by).

Fig. 3. Child painting the TV Companion (left) and customized RFID tokens (right)

4 Study

We conducted a situated field study of the TV Companion with three families in order to gain insights regarding usability, user acceptance and in particular regarding the aforementioned design concerns. The study was preceded by a pre-study and included a workshop with initial interviews, a 10 day study phase and a final interview.

4.1 Study Setup

To investigate the TV Companion concept and to get first feedback we conducted a pre-study workshop with a family and their three children (see Table 1) using an early prototype. During this workshop the children painted the device and the tokens and the family was interviewed. Based on both their enthusiasm for the idea and feedback, we then built three identical TV Companions and deployed them in the homes of three families, each of whom had a four year old child, and two also had a second younger child. All families described themselves as regular TV consumers. The pre-study participants were recruited from the extended social network of the authors and the families in the study were recruited via a local kindergarten. The study lasted for 10 days and consisted of an introductory in-home workshop and interview, the study phase in which the TV Companion was used by the families, and a final interview.

Table 1. Summary of families included in the study

Family	No. of children	Age of children	Phase
Family 1	3	4, 6, 11	Pre-study
Family 2	1	4	Study
Family 3	2	1, 4	Study
Family 4	2	2,4	Study

During the initial workshop, the function of the box was explained to the parents and the child. An initial interview regarding the TV viewing behavior and their first impressions of the TV Companion was conducted. The kids then had a chance to

customize their TV Companion by painting it and the tokens with acrylic colors. The families then chose a suitable spot for the TV Companion in the living room and tested, with the support of the researcher, whether the device functioned properly. This included in particular to see if the device was able to communicate with the TV set and could turn it off at the end of the viewing time. The families were told to use the box in the way that suited them best, leaving aspects like the way they distributed the tokens to their children up to them, and without any obligation to use it at all.

In order to get insights about the participants experience with the TV companion during the study, they were given a diary with the request to take a couple of minutes to fill it out after the children would have gone to bed. The diary included structured questions regarding the amount of time the child watched TV, if he/she used the box, whether he/she made any comments about the TV Companion and room for open ended statements and observations. To complement the diary, participants were equipped with a digital camera and instructed to occasionally use it to take a picture related to their TV viewing.

Fig. 4. Child putting RFID token on the TV Companion (left). Storing *her* tokens in a little wooden drawer box.

At the end of the study the parents and the children were interviewed again. This semi-structured interview also took into account any issues that came up from an analysis of the diaries and included questions about their TV viewing during the 10 days of the study, their interaction with the TV Companion and whether they had any suggestions regarding its design. The adult study participants were not financially remunerated but the children received a little toy as a token of appreciation for their efforts. The data of the interview recordings and notes, the diaries and the photos taken were analyzed qualitatively to get an understanding of the use of the device, the usage patterns that occurred and the experiences of the families with TV Companion.

4.2 Results and Discussion

During the initial interviews, the parents (Family 1,3,4) said that they had the feeling their children were watching too much television and they were enthusiastic about the idea of the TV Companion. They thought it could help them to reduce the TV consumption of their children. While the children agreed that the TV companion would be fun to use and that watching too much TV is not good - "your eyes get rectangular and start to hurt" - as one of them pointed out, some of them (Family 1) disagreed with their parents that they were actually watching too much TV themselves.

Overall, the participants reported the TV Companion was easy to use; their children had no problems interacting with the device. Family 3 and 4 stated that they used the device at least once per day, whereas Family 2 used it several times per week in accordance to their established TV viewing habits. The color feedback changing from green to red and the voice prompts were considered clear and understandable for the children. On a technical level, the device worked well with the TV set of each family and was able to turn it off after the time had run out.

Two of the families (3,4) pointed out the fact that after the box has announced that the time has run out it is not possible anymore to add another token and just continue watching without the TV being turned off by the box. While the parents saw this as a good thing, since it interrupted the TV watching process and sent a clear signal to the children that they used up this token, the children would have liked the opportunity to just add another token and continue watching if they wanted to and if they still had a token left. In the interview after the study, Family 4 particularly emphasized that they liked the simple and reduced design, and especially the intuitive and friendly voice of the TV Companion's audio feedback.

The study indicates that the children enjoyed the process of customizing the box. They were eager to paint the wooden box of the device and also made use of the possibility to decorate the tokens (see Figure 3). The possibility to customize the device by painting it, which was based on the design concern *autonomy and free choice* was enjoyed by all of the children. One family explained that this kind of appropriation lead to a sense of ownership of the box in their daughter, as we had hoped. This resulted in her liking to use the TV Companion and also in using it regularly. She also treated the tokens like her other valuable possessions such as (toy) jewelry and placed them in a special drawer (see Figure 4).

Family 4 reported that they also use the computer to let their child watch DVDs, and even though the computer screen is not compatible with the TV Companion, they symbolically let their child use a token with the device also for this activity since they wanted to show him that this is part of the time budget. Hence, in this instance the user could *autonomously* adapt the systems to his ends.

Family 2 pointed out in the initial interview that in their education of their child they do not see television watching as an incentive and/or punishment system, but the design of the TV Companion implicitly contains these values. In their view this is evident by the systems use of "money-like" tokens to symbolize the time allotted to the child to "spend". They already didn't let their child watch TV at will but carefully selected the shows she was allowed to watch. Nevertheless, they found a way to use

the system to support their own approach to managing TV viewing by giving single tokens to their child before each selected show, thus making sure that she would only watch that show and not continue watching after this specific program has ended. This manner of adapting the TV Companion to their own ideas supports the notion of ends-based tailoring brought forward in the design concern regarding *autonomy*. As the diary and the interview indicates, in this tailored approach it was mostly the parents who initiated the use of the box, which does resonate with way this family used TV prior to the introduction of the box.

Family 3 and 4 both let their children watch TV more freely and shared the concern that by doing so their children were watching too much television. They also reported that their children would turn the TV on, watch something for a while and then start doing other activities while the TV would remain on in the background, thus creating a nuisance for the other family members. Both expressed the hope that the TV Companion would positively influence their children's habit of leaving the TV on even when they were not watching. Coming back to this issue in the final interview both families reported that the TV Companion changed this behavior significantly since the children did not want to waste any tokens and only turned the TV on and invested a token when they actually wanted to watch a program.

Family 2 and 3 allocated tokens to their children on the basis of each TV viewing session, whereas family 4 gave their child a token budget for the entire week. They reduced the time to one day during the study to decrease excessive TV consumption on single days caused by the child spending a large portion of the tokens at once. After learning that spending all the tokens at once really meant no more TV for that day, their child used the tokens to specifically watch only the shows she was interested in, leaving the TV turned off otherwise. This hints at the possibility that the TV Companion might be able to foster *reflection* as proposed in the second design concern.

The "wiggle room" provided by the systems based on the design concern *coercion versus reflection* was used by the child of family 4. He knew that the system was flexible, and while he always used it with his mom, when his dad, who has a busy schedule as ice hockey trainer, came home and they decided to watch a game together, he choose not to spend any tokens for this special occasion.

The fact that the TV Companion did not use *surveillance* to track the amount of TV consumed by the children was appreciated by all families participating in the field study. One the one hand they thought it was sufficient to calculate the TV viewing time based on the tokens used. On the other hand they even stated that they are opposed to the possibility of sharing this information e.g. via a social network, since this was a private family matter. Family 3 pointed out that this might even lead to an unwanted competition about who was a better parent based on the TV behavior of their children.

5 Conclusion and Future Work

This paper explored three design concerns of PT technology systems: *autonomy and free choice*, *coercion versus reflection* and *surveillance and privacy* and their interpretation in a prototype system called the TV Companion aimed to help parents and

children manage their own TV viewing habits. Through a pre-study followed by three in-home explorative field trials, we found evidence that participants were able to adapt and use the designed features of the TV Companion in different ways to suit their own ends.

There is certainly the need to support the initial findings by further evidence and promising feedback from both the parents and the children encourages us to dig deeper into the concept of persuasive technologies with built-in "wiggle room" and opportunities for reflection, designed for the autonomous reflective user. Hence, it is planned to collect additional data from more participants in long-term studies, enabling us to make inferences about induced behavior change.

In the upcoming studies special light will also be shed upon customization in PT systems. We have already observed positive results with the TV Companion concerning customization and appropriation. Therefore, we aim to extend this concept beyond the pure painting of the device and RFID tokens. One design opportunity, e.g., is the recording of own voice messages. This leads to questions such as, *whose* voice should be recorded (the kids, the parents or maybe a stranger?), *what* statements should be recorded (content and tone of voice) and *when* should voice messages be played. Can these additional degrees of customization and autonomy provide further support in encouraging positive behavior?

Finally, we have observed an interesting side effect of the TV Companion. Since the device automatically shut down the TV after the defined period of time, there were no unattended TVs running in the background – a common habit at the home of many people. Therefore, it might be well worth to address the design concept of a persuasive technology with open means as well as open ends in other fields than positive health behavior change, for instance in the field of sustainability.

Acknowledgements. We would like to thank all of our participants for testing the prototype and for their valued feedback. Our very special thanks are due to Dorothee Köstlin for her thought-provoking impulses and contribution to the prototype.

References

1. Obermair, C., Reitberger, W., Meschtscherjakov, A., Lankes, M., Tscheligi, M.: perFrames: Persuasive Picture Frames for Proper Posture. In: Oinas-Kukkonen, H., Hasle, P., Harjumaa, M., Segerståhl, K., Øhrstrøm, P. (eds.) PERSUASIVE 2008. LNCS, vol. 5033, pp. 128–139. Springer, Heidelberg (2008)
2. Lin, J.J., Mamykina, L., Lindtner, S., Delajoux, G., Strub, H.B.: Fish'n'Steps: Encouraging Physical Activity with an Interactive Computer Game. In: Dourish, P., Friday, A. (eds.) UbiComp 2006. LNCS, vol. 4206, pp. 261–278. Springer, Heidelberg (2006)
3. Consolvo, S., McDonald, D.W., Toscos, T., Chen, M.Y., Froehlich, J., Harrison, B., Klasnja, P., LaMarca, A., LeGrand, L., Libby, R., Smith, I., Landay, J.A.: Activity sensing in the wild: a field trial of ubifit garden. In: Proceedings CHI 2008, pp. 1797–1806. ACM, Florence (2008)
4. Gustafsson, A., Katzeff, C., Bang, M.: Evaluation of a pervasive game for domestic energy engagement among teenagers. Computers in Entertainment 7, 1 (2009)

5. Veerman, J.L., Healy, G.N., Cobiac, L.J., Vos, T., Winkler, E.A.H., Owen, N., Dunstan, D.W.: Television viewing time and reduced life expectancy: a life table analysis. British Journal of Sports Medicine (2011)
6. Lillard, A.S., Peterson, J.: The Immediate Impact of Different Types of Television on Young Children's Executive Function. Pediatrics (2011)
7. Education, C.o.P.: Children, Adolescents, and Television. Pediatrics 107, 423–426 (2001)
8. Hofkirchner, W., Tscheligi, M., Bichler, R., Reitberger, W.: Ambient Persuasion for the Good Society. IRIE 8, 42–46 (2007)
9. DiSalvo, C., Sengers, P., Brynjarsdóttir, H.N.: Mapping the landscape of sustainable HCI. In: Proceedings CHI 2010, pp. 1975–1984. ACM, Atlanta (2010)
10. Friedman, B., Kahn, P., Borning, A.: Chapter 16: Value Sensitive Design and Information Systems. In: Human-Computer Interaction and Management Information Systems: Applications, pp. 348–372 (2006)
11. Fogg, B.J.: Persuasive Technology: Using Computers to Change What We Think and Do. Morgan Kaufmann Publishers Inc. (2003)
12. Oinas-Kukkonen, H., Harjumaa, M.: Persuasive Systems Design: Key Issues, Process Model, and System Features. Communications of the Association for Information Systems 24 (2009)
13. Kaptein, M., Duplinsky, S., Markopoulos, P.: Means based adaptive persuasive systems. In: Proceedings of the 2011 Annual Conference on Human Factors in Computing Systems, pp. 335–344. ACM, Vancouver (2011)
14. Purpura, S., Schwanda, V., Williams, K., Stubler, W., Sengers, P.: Fit4life: the design of a persuasive technology promoting healthy behavior and ideal weight. In: Proceedings of the 2011 Annual Conference on Human Factors in Computing Systems, pp. 423–432. ACM, Vancouver (2011)
15. Reitberger, W., Ploderer, B., Obermair, C., Tscheligi, M.: The PerCues Framework and Its Application for Sustainable Mobility. In: de Kort, Y.A.W., IJsselsteijn, W.A., Midden, C., Eggen, B., Fogg, B.J. (eds.) PERSUASIVE 2007. LNCS, vol. 4744, pp. 92–95. Springer, Heidelberg (2007)
16. Leth Jespersen, J., Albrechtslund, A., Øhrstrøm, P., Hasle, P., Albretsen, J.: Surveillance, Persuasion, and Panopticon. In: de Kort, Y.A.W., IJsselsteijn, W.A., Midden, C., Eggen, B., Fogg, B.J. (eds.) PERSUASIVE 2007. LNCS, vol. 4744, pp. 109–120. Springer, Heidelberg (2007)
17. Foucault, M.: Discipline and punish: the birth of the prison. Vintage Books, New York (1995)

Bridging the Gap between the Home and the Lab: A Qualitative Study of Acceptance of an Avatar Feedback System

Peter A.M. Ruijten, Yvonne A.W. de Kort, and Petr Kosnar

Eindhoven University of Technology, The Netherlands

Abstract. The current study provides a first step in the design and development of a persuasive agent in the natural context of the household. We developed two persuasive probe studies: one paper-based probe and one email-based probe on the use, experience, and effectiveness of persuasive agents. Participants had used these prototypes for a week, after which their experiences were explored in depth interviews and a focus group. Results indicated that a persuasive agent in the household is experienced as fairly pleasant, but important issues need to be solved before it can effectively influence behavior.

Keywords: Persuasive Technology, Persuasive Probes, Sustainable behavior.

1 Introduction

To date not much is known about the aspects that influence agent acceptance and effectiveness after repeated exposure in naturalistic settings. To be able to successfully employ an embodied artificial agent in the home environment, interactions with, and effects and perceptions of these types of applications should therefore be studied in their natural habitat and with a more longitudinal perspective. The current study provides a first step in this direction, applied in the context of energy conservation behavior.

Research has demonstrated that so-called prompts are most effective there and then where (negative) behavior is performed [e.g., 1]. Reminding users to use less water while they are changing thermostat settings or to switch off the television while brushing their teeth will probably be ineffective. Moreover, the optimal tone of persuasive messages (e.g. directive or friendly) may also differ depending on whether users experience them a few times during an hour of experimentation, or repeatedly throughout the day, for months in a row. In the current study, this issue will be addressed by tapping people's experiences with both the content of the messages and the avatars' expressed emotions.

Investigating these aspects is by definition impossible in controlled lab environments, but should instead be performed in the natural context of the envisaged application. Technology probes [2, 3] are a research methodology designed to gather information about possible end users of systems in real life, testing a system in a field setting, and inspiring the user to assist in the design phase. A technology probe generally is an instrument that is developed to find out things about the non-existing

M. Bang and E.L. Ragnemalm (Eds.): PERSUASIVE 2012, LNCS 7284, pp. 251–255, 2012.

technology [3]. In the context of our study, we did not test a functional application, but instead created non-interactive paper prototypes. Therefore we refer to our methods as persuasive probes instead of technology probes, whilst pursuing the same goals. The persuasive agent that we will probe in this study is one that provides direct feedback about people's energy-efficient behavior. To examine these and other issues, paper prototypes and non-interactive digital prototypes were employed in a persuasive probe type study to investigate the experience, use, and effectiveness of a persuasive agent.

2 Methods

To investigate the persuasive potential of an avatar feedback system in the home environment, we tested two persuasive probes. The first one was a paper prototype-based persuasive probe study on the use, experience, and effectiveness of persuasive agents in real life *physical* environments (paper probe) and the second one was an email-based persuasive probe study on the use, experience, and effectiveness of persuasive agents in *digital* environments (email probe).

2.1 Participants and Design

In the paper probe study (PPS), three households (CPL1-3) were asked to participate. After six days, an interview (INT) was conducted to discuss people's experience with the simulated system. Each household was given a voucher of €25 for their participation.

In the email probe study (EPS), six participants (SUB1-6, $M_{age} = 25.0$; $SD_{age} = 2.61$) were asked to participate. They were sent three, six or a random number between one and six of emails per day. This study also lasted for six days, after which we performed a focus group session (FG). Each participant was paid €15 for their collaboration in this study.

2.2 Materials and Procedure

Paper-Prototype Probe Study. Three households were given a set of 12 stickers with a size of 10.5 × 12 cm, each sticker containing a picture of an avatar displaying an emotion (neutral, happy, sad, angry or surprised) and a related message. Furthermore, they were provided with small paper notebooks and pens to place close to each sticker. On these they could write all thoughts and suggestions they had regarding this specific sticker at any time during the study. The messages were chosen to describe situations for various energy demanding behaviors or to give suggestions about advantages of such conservation behavior. A detailed list of these descriptions is provided in Appendix I.

Email Probe Study. The content of the emails was similar to the content of the stickers in the sticker study. Each email message contained an image of an avatar showing an emotional expression combined with a message related to energy consumption and environment-friendly behavior. Participants were given the opportunity to reply to the emails if they wanted. After a total of six days, participants reflected on their experience either in a focus group or individual interview. Because we used the same discussion themes in both studies, data from all interviews and the focus group were collapsed.

3 Results

3.1 Messages

Most of the participants in both studies reported that they changed their behavior in accordance with the messages they received. One participant became aware of the fact that using the television for background noise uses quite a lot of energy, and changed his behavior accordingly:

> *"One message was about turning on your radio instead of TV for hearing some music or something like that. […] So you [can] better turn on the radio, and I did it. On that perspective it was very useful."* [SUB4, EPS, FG]

3.2 Avatars

Some participants in the paper probe study reported that the mere presence of the avatars in their homes caused them to pay more attention to their energy efficient behavior. Between the participants, there was some debate about how the avatar should look. Some people would like it to be as humanlike as possible, whereas others would prefer a more cartoonlike character. Some people stated that the avatar should look like somebody they know, like a parent, which gives the avatar some authority.

3.3 Interaction

For most people it is important that the given feedback is dynamic and context aware. That is, the feedback that is provided by the system should assist people in their current activities, both in message context and in the avatar's emotional expression. In both studies, people perceived the feedback coming from the avatar as pleasant. After a while participants did get used to the avatar's presence. This made people feel the avatar as becoming part of their family in one of the households:

> *"She was included in the family rather quickly. She was just part of it. [Interviewer] What name did you give her? [Participant] Marie."* [CPL1, PPS, INT]

4 Discussion

The current study was designed to investigate the use, experience and effectiveness of persuasive agents in both real life *physical* environments (paper-prototype probe) and in *digital* environments (email probe). Many participants perceived the messages as useful, helpful and they reported that it influenced their behavior.

Considering the tone of the persuasive messages, people experienced the positive emotions as more helpful than the negative ones. This does not match results from earlier research [e.g., 4, 5]. We argue that in future applications, it is important that the given feedback fits the context of the situation, so information about what the user is actually doing at that moment is essential to implement. The mere presence of the agent probes in people's lives caused them to pay attention to their energy related behavior. However, people felt as if the provided probes were unable to adapt to their personal needs, which could be solved by adding some personalization aspects in the system, such as the possibility to select the avatar's appearance.

The results from this study provide a first view on how an interactive, persuasive system will be experienced, as well as some interesting insights into people's preferences for such a system to be effective. These findings are valid as information source for the development of further research, but no general statements should be drawn from them because of the low number of participants that used the probes. We believe that an agent in the home can help and assist people to save energy in the future, but some of these aspects should be solved before people will actually accept this new technology as part of their daily lives.

Acknowledgements. We would like to thank Andrei Antal, Marco Knops, and Remy Kaufman for their assistance and collaboration in this research project.

References

1. Geller, E.S., Witmer, J.F., Tuso, M.E.: Environmental interventions for litter control. Journal of Applied Psychology 62, 344–351 (1977)
2. Graham, C., Rouncefield, M., Gibbs, M., Vetere, F., Cheverst, C.: How probes work. In: Proceedings of OzCHI 2007, pp. 29–37 (2007)
3. Hutchinson, H., et al.: Technology probes: Inspiring design for and with families. In: Proceedings of CHI 2003, pp. 17–24 (2003)
4. Ham, J., Midden, C.: A robot that says bad! Using negative and positive social feedback from a robotic agent to save energy. In: Proceedings of the 4th ACM/IEEE International Conference on Human-Robot Interaction, pp. 265–266 (2009)
5. Midden, C., Ham, J.: Using negative and positive social feedback from a robotic agent to save energy. In: Proceedings of the 4th International Conference on Persuasive Technology, Claremont, CA, USA, April 26-29, p. 12 (2009)

Appendix I

A list of all descriptions, translated from Dutch. Each set of twelve pictures contained the messages presented below. In the email probe study, similar but shorter messages were used.

1. Take the environment into account by keeping an eye on how long you stand in the shower. In general, people use 30% more water in taking a bath compared with taking a shower!
2. Using the standby and sleep mode on your computer can save considerable amounts of energy. Most people let their computer switch to the sleep mode after fifteen minutes. This is still unnecessarily long.
3. Consider turning down the heating a bit when you are not cold. Did you know that around 7% of energy can be saved in your heating costs by turning down the heating by just one degree?
4. People who turn off their electrical devices when they leave their homes make me happy. Do not use energy unnecessarily by having the computer in sleep mode all day.
5. I am happy that so many households change their light bulbs by CFL's. You should do this as well! These lamps save up to 75% of current!
6. I am happy that you separate your waste. Most people these days are not consciously engaged in correctly separating their waste.
7. I think it is a shame that you leave your television on standby unnecessarily. Leaving on electrical devices for no good reason is a huge waste of current.
8. It is wrong to leave your stereo on when you leave the house. It is a waste of energy when people leave their electrical devices on when nobody is close to them.
9. Try washing your clothes at 30 degrees. Your clothes will still become clean, and you will save energy!
10. Are you familiar with the energy labels on electrical devices? Devices with a green energy label are environmentally friendly, and will save you money in the long term as well!
11. I think it is a waste of water when people leave the tap open while they brush their teeth. Brushing your teeth takes about two minutes. So a large amount of water will be wasted when you leave the tap open.
12. Why don't you turn off the light when you leave the room? Thinking about the amount of energy that is wasted in this way makes me angry.

Each of the three avatars that were used displayed five different emotions: neutral, happy, sad, surprised and angry. In Figure 1, the neutral expressions are presented.

Fig. 1. The avatars used in both probe studies, presented with the neutral expression

Less Fizzy Drinks: A Multi-method Study of Persuasive Reminders

Sitwat Langrial and Harri Oinas-Kukkonen

University of Oulu, Department of Information Processing Science,
Rakentajantie 3, FI-90570, Oulu, Finland
{sitwat.langrial,harri.oinas-kukkonen}@oulu.fi

Abstract. In this paper, we present initial outcomes from our multi-method study that explored the impact of reminders on effectiveness of information systems that aim to facilitate behaviour change. Our study explicitly focused on reminders as a key persuasive software feature that should be employed to support behaviour change. We designed a prototype aiming to encourage people to gradually reduce soda/fizzy drinks consumption. A multi-method 14-day pilot study was conducted, composed of statistical analysis followed by a qualitative focus group. Statistical analysis shows that participants consumed less fizzy drinks in the second half of the study, and feedback from the focus group study indicates that reminders successfully persuaded participants to keep a log of their fizzy drink consumption. Our study supports the claim that persuasive reminders have extraordinary potential for helping people change their behaviours.

Keywords: Behaviour change, persuasive systems design, persuasive features, reminders.

1 Introduction

Consumption of fizzy (soft) drinks such as soda and cola instead of water has become a common habit. Fizzy drinks are high in sugar, calories and addictive ingredients and possess little or no dietary value. Such drinks contain phosphoric acid, sugar, caffeine and other harmful ingredients. Soft drinks (both fizzy and soda) result in weight gain, which in turn leads to an increased risk of diabetes.[1] Another side effect of such drinks is weakening of bones and higher risk of osteoporosis.[2] Despite all of this publicly available information, it is evident that people do not view fizzy drinks consumption as habit with negative health effects. We developed a web-based persuasive prototype (system) that aimed to support people in reducing fizzy drinks consumption.

[1] www.emedexpert.com/tips/soft-drinks.shtml. Accessed on 28-12-2011.

[2] www.nhs.uk/news/2009/05May/Pages/ColaSapsMuscleStrength.aspx. Accessed on 29-12-2011.

M. Bang and E.L. Ragnemalm (Eds.): PERSUASIVE 2012, LNCS 7284, pp. 256–261, 2012.

Persuasive systems can be defined as "computerized software or information systems designed to change or shape attitudes or behaviours or both without using coercion or deception" [1]. Several factors may hinder performance of the desired behaviour. These may be linked to motivation, ability, resources, time, or reminders [2], which belong to the key persuasive software features in behaviour change technologies [3]. Reminders been employed in scores of studies aiming to promote healthy living; for instance, Fry and Neff [4] argue that intervallic reminders are efficient in behaviour change technologies, while Fjeldsoe et al. [5] argue that reminders delivered via short message service have a positive effect on behaviour change. Bickmore et al. [6] argue that an insistent and obtrusive reminder might well be effective in the short run, but the effect reduces significantly over time. On the other hand, if the reminders are polite, there is a likelihood that little or no compliance would occur in the early stages; however, users are more likely to use the device over the long run. Walji and Zhang [7] employed human-centred design principles to develop persuasive reminders that were intended to encourage people (patients) to be present at their allocated appointment. They argue that user-centred design approach proved to be valuable and helped them make several modifications to their appointment reminder system.

This paper seeks to investigate the impact of reminders in persuasive systems. The aim is to analyse in what way and to what extent reminders help users perform desired behaviours. The paper presents the findings of a 14-day pilot study followed by outcomes of a focus group that was conducted at the University of Oulu, Finland.

2 Theoretical Viewpoint

To design and develop our prototype by employing a framework that outlines fundamental persuasive features, we expended the Persuasive Systems Design (PSD) Model put forward by Oinas-Kukkonen and Harjumaa [3]. The PSD Model encourages thorough analysis of the context of a persuasive attempt, i.e. the intent, the event and the persuasive approach adopted. Further, it signifies the importance of identifying key contents and software functionalities for designing and developing persuasive systems. It supports a comprehensive consideration of persuasive features during the design process of a persuasive system, and provides a foundation for a feature-by-feature evaluation of a persuasive system. The PSD Model outlines four distinct categories of persuasive features. These categories are: 1) Primary Task Support (aimed at supporting users' primary tasks), 2) Dialogue Support (aimed at helping users reach their goals), 3) System Credibility (aimed at improving system credibility) and 4) Social Support (aimed at enhancing users' motivation through social influence).

3 Research Setting

3.1 Participants and Recruitment

The participation was required to be on a volunteer basis and consisted only of those who consumed one or more fizzy drinks on an average day and were willing to reduce

fizzy drinks consumption. For the purpose of recruitment, we sent out e-mails to the Department of Information Processing Science at the University of Oulu, Finland. In addition, we visited some classrooms and explained the study. Following this, 29 participants were enrolled for the pilot study. Eight females (27.6%) and 21 males (78.4%) with a minimum age of 22 and maximum age of 34 years volunteered to participate. To compensate their participation, a lucky draw (€50 Amazon Gift Voucher) was organised for one winner at the end of the study.

3.2 About the Less Fizzy Drinks System

All the registered participants were sent a "welcome" e-mail along with a URL to the system's registration page. Using their e-mails as unique user ID, they were required to fill out a form providing information about their age, gender, education, profession, weight, telephone number and average consumption of fizzy drinks. An activation link was auto-generated and sent to the participant's e-mail upon completion of registration. On the first day of their interaction with the prototype, users recorded their average fizzy drinks consumption as well as their current weight. For the remaining days, they received a daily reminder through e-mail at 9 a.m. prompting them to record their fizzy drinks consumption for the day before. This was to ensure that they did not miss out on any fizzy drinks intake that they might consume after having recorded their daily consumption. Users were also required to record their weight two times during the 14-day interaction with the system, i.e. day 7 and day 14. Two separate reminders were sent out to the participants regarding this, one on day 7 and the second on day 14. Similar to the daily reminders, the weight update reminders were e-mail-based, promoting the participants to update their weight. Each time a user recorded fizzy drinks consumption, the system responded with immediate feedback. The content of the feedback depended on the amount of recorded fizzy drinks consumption. For example, when a participant recorded a reduced number of drinks, praise was instantly prompted. For example, *"Well done. You are successfully reducing your fizzy drinks intake."* Or when a participant recorded the same amount of fizzy drinks intake as the day before, the system prompted, *"Your fizzy drinks intake is the same as yesterday."*

3.3 Post-Study Focus Group

Following the 14-day pilot study, users were invited to be part of a focus group. Thirteen (n=13) volunteered to participate. A 30-minute focus group was conducted during which they were served refreshments. With their consent, the discussion was audio recorded.

4 Results

4.1 Statistical Analysis

We wanted to examine whether the users adhered to a routine of updating their drink consumption record upon receiving daily reminders: It was our hypothesis that the

effect of the reminders would be such that the users would keep updating their drink consumption log. Further, we wanted to detect any change in the amount of fizzy drink consumption during the pilot study; detecting even a slight change in consumption was considered as an indicative factor for a possible behaviour change in future work. The results in terms of task adherence relating to reminders (to update drink consumption record) were staggering, as all the participants (N=29; 100%) updated their records for the entire study period. We acknowledge the small sample size and short study duration, in which users were not required to perform time-consuming tasks. However, the fact that all the users responded positively to the reminders is a promising finding and cannot be overlooked.

For the purpose of descriptive statistical analysis, we decided to split the data into two halves, i.e. average drink consumption for the first seven days and the last seven days. A paired-sample T-test was conducted to compare the average drink consumption in Week 1 (Condition 1) and Week 2 (Condition 2). Results from T-tests indicate that there was a significant difference between mean of average soda drinks consumption for Week 1 (M=4.89, SD=4.64) and Week 2 (M=2.27, SD=2.98). The results from the Paired Differences between Week 1 and 2 with a confidence interval of 95% and Sig. (2-tailed) value of (p=0.003) suggest that persuasive reminders helped users conform to the subtask of recording their intake.

We interpret these results in such a manner that there was a positive effect of our system on the participants. Specifically, after users' interaction with the prototype for the first week, there was a noticeable reduction in soda drinks consumption. However, given the limited scope of the study, it cannot be claimed that the reduction of fizzy drinks consumption was as a consequence of the persuasive reminders only. However, the significant reduction in mean value (M=2.62) encourages us to argue that further improvement to the prototype and a field study may reveal more in-depth results.

4.2 Focus Group Results

A focus group was conducted to gather users' insight and feedback on about the effectiveness of the reminders and overall perceived persuasiveness of the software features. We opted for a focus group because they are similar in nature to individual interviews [7]; however, the concurrent participation stimulates a more in depth discussion in the presence of a facilitator. Participants of the focus group were asked about their evaluation of reminders that they received during their interaction with the system. It was stipulated that they should keep in mind the content, frequency and timing of the reminders while commenting. All the (n=13,100%) participants highly approved the reminders and gave encouraging feedback. There were some worthy comments about the timing and frequency of the reminders during the discussion. The general consensus amongst the participants was that in such a study, one reminder a day is ample. Participants were then asked to give their opinion and feedback regarding the timing of the reminders. A high majority of the participants (n=12) expressed positive remarks about the timing of the reminders. In response to whether

the reminders were useful, all the participants gave positive remarks about the effectiveness of the reminders. Some of the representative comments (C) are listed below:

C1: "I would say that the reminders were well (good) because when I check my e-mail in the morning, it was the first thing that I did. It (the procedure) was really simple. You go in there (to the e-mail) and one click will take you to the system. It would take fifteen seconds to complete the task. So they (reminders) worked well"." (Male, Graduate student)
C2: "I think reminders were good. For the first two times (days), I used the reminders. But from the third day, I recorded my (fizzy drinks) consumption unconsciously." (Male, Web designer)
C3: "The system helped me a lot because I am more driven by reminders." (Male, UI designer)
C4: "I go to the system after receiving the reminders because now I can trust that I am being reminded and I do not need to remember (by myself)." (Female, Graduate student)

5 Discussion

The statistical analysis from the study indicates that the reminders were successful in motivating users to regularly update their fizzy drink consumption. Further, a significant reduction in consumption is observed. Participants of the focus group showed a prodigious approval of the reminders' persuadability. They were perceived as simple, on time and facilitating use of the system. Interestingly, none of the participants reported being irritated by the frequent (daily) reminders. The prototype proved to be robust enough for the pilot study. Based on the data analysis and feedback received from the users, we intend to make amendments to the prototype both in terms of usability and additional persuasive features.

Our study was prone to several limitations. First, the sample size was not large enough to generalise results. Second, the subjects of the study do not represent the broader population, especially those who consume fizzy drinks heavily. Further, for the pilot stage we did not have a control group that had the same intervention without reminders. Nevertheless, the mixed-method approach provided an in-depth insight to our study. We therefore argue that the discipline of persuasive systems has reached a stage where multi-method studies could lead to better designing and implementation. User participation in the design process can enhance the effectiveness of persuasive systems strategies and features.

6 Conclusions

We investigated the scope of persuasive reminders to encourage users to reduce fizzy drink consumption. It was found that reminders served the purpose of persuading people to record their daily fizzy drink intake. All of the users adhered to a routine of

recording their fizzy drinks consumption for the entire duration of the study. We argue that this is a promising finding.

The focus group discussion highlighted useful ideas that will be utilised to further refine the prototype. Various themes emerged from the discussion in the focus group that covered different aspects of designing persuasive systems. From the findings (quantitative and qualitative) of the study, we can safely argue that reminders are an essential part of any given persuasive system. Our study presents a blend of quantitative and qualitative methodologies. It offers openings for further research where valuables insights could be achieved beyond quantitative data.

Acknowledgements. We thank Dr. Jouni Markkula for his help with statistical analysis. We appreciate Payam Hossaini for his contribution towards building the prototype. This study was supported by SalWe Research Program for Mind and Body (Tekes – the Finnish Funding Agency for Technology and Innovation grant 1104/10).

References

1. Oinas-Kukkonen, H., Hasle, P., Harjumaa, M., Segerståhl, K., Øhrstrøm, P. (eds.): PERSUASIVE 2008. LNCS, vol. 5033. Springer, Heidelberg (2008)
2. Fogg, B.J.: Creating Persuasive Technology: An Eight Step Process. In: Persuasive 2009, California, USA (2009)
3. Oinas-Kukkonen, H., Harjumaa, M.: Persuasive Systems Design: Key Issues, Process Model, and System Features. Communications of the Association for Information Systems 24(1), Article 28 (2009)
4. Fry, J.P., Neff, R.A.: Periodic Prompts and Reminders in health promotion and Health Behavior Interventions: Systematic Review. J. Med. Internet Res. 11(2), e16 (2009)
5. Fjeldsoe, B.S., Marshall, A.L., Miller, D.: Behavior Change Interventions Delivered by Mobile Telephone Short-Message Service. Am. J. Prev. Med. 36(2) (2009)
6. Bickmore, T., Mauer, D., Crespo, F., Brown, T.: Persuasion, Task Interruption and Health Regimen Adherence. In: de Kort, Y.A.W., IJsselsteijn, W.A., Midden, C., Eggen, B., Fogg, B.J. (eds.) PERSUASIVE 2007. LNCS, vol. 4744, pp. 1–11. Springer, Heidelberg (2007)
7. Walji, F.M., Zhang, J.: Human-centered Design of Persuasive Appointment Reminders. In: Proceedings of the 41st Hawaii International Conference on Systems Science (2008)

Plotting to Persuade – Exploring the Theoretical Cross Field between Persuasion and Learning

Sandra Burri Gram-Hansen, Henrik Schärfe, and Jens Vilhelm Dinesen

Department of Communication and Psychology, Aalborg University
{burri,sharfe}@hum.aau.dk, dines10@student.aau.dk

Abstract. This paper presents a few of the initial reflections related to ongoing research on the notion of Persuasive Learning. In addition the paper briefly comments upon some of the benefits and challenges related to the cross field between persuasive technologies and learning objects, and between persuasion and learning in general.

Keywords: Persuasive Design, Pedagogy, Kairos, Didactics, e-PLOT.

1 Introduction

The literature on both learning designs and persuasion are vast, yet these areas have not previously been combined in a systematic manner. The initial steps taken towards defining and exploring the areas in which persuasion and learning may complement each other, are based on the belief that both persuasion and learning are highly contextual phenomena where intentions of the designs and the users and the negotiation of intentions are equally important to the applied technology. Based on current research performed within the EU-funded research project Euro-PLOT, this paper presents the primary theoretical foundation for the development of persuasive learning designs, and discusses some of the conflicts and challenges related to exploring and defining the cross field between persuasion, pedagogy and didactics.

2 Persuasive Design in a Humanistic Perspective

The notion that the theoretical and methodological aspects of Persuasive Technology might benefit from applying classic humanistic traditions such as rhetoric, logic and ethics to its development, has previously been argued by various researches [1, 2]. In the process of defining a theoretical foundation for the development of Persuasive Learning Designs (PLDs), these different traditions have been brought together in the attempt to define a foundation which not only clarifies immediate overlaps between persuasion and learning, but which also presents the notion of Kairos as a core concept within persuasive design.

One area which has been of particular interest in the exploration of persuasive learning, is the notion of persuasive intentions. The designer will most often have a

M. Bang and E.L. Ragnemalm (Eds.): PERSUASIVE 2012, LNCS 7284, pp. 262–267, 2012.

specific intention with the design of a technology, but this intention is often more complex than "wanting to motivate people to quit smoking", and the user's intention towards applying a specific technology, must to some extent be motivated by exogenous factors. As a result, the notion of intention in relation to Persuasive Design also calls for an adjusted and more nuanced perspective [3].

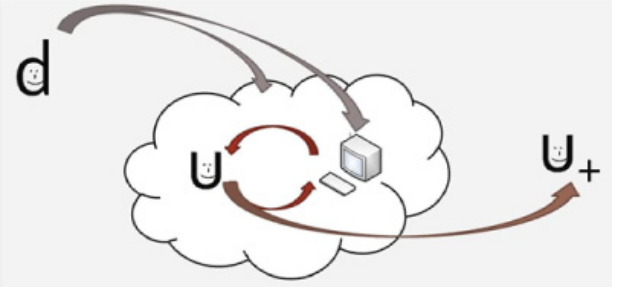


Fig. 1. The intention of a technological design concerns both the technology and the use context

Figure 1 illustrates that the designer's intention concern both the technology, and the context in which the technology is to be applied. Once applied within the intended context, the user and the technology reciprocally influence each other, resulting in an alteration of the context and in the user being persuaded to change attitude or behaviour. The intention remains a human characteristic, whilst the technical capacity of the technology serves to support the fulfilment of the persuasive intention. In order for a persuasive intention to be successfully met, there must be an appropriate balance between the endogenous intention of the designer and that which motivates the user to apply the technology. Designers are limited to conceiving the intended use of a technology, with no means to assure that the practical use will correspond. Once a technology is put to use, its employment as well as the users perception of the endogenous intention, is influenced by the context in which it is applied. When applied within the appropriate context, a technology may prove itself to be a highly efficient persuader, but if applied without consideration of the context, the Persuasive Design could fail to fulfil its persuasive goal, or be considered unethical [3].

The rhetorical notion of Kairos is often described as timing, or the ability to perform the appropriate action at the right time and in the right place. In term of appropriate, the performed action is required to be not only effective but also ethical.

The concept sums up the principle that any rhetorical approach is based upon the specific situation, and that comprehension of the context as such is one of the most vital resources when deciding upon rhetorical means to apply to a given argument [4] Hansen specifies that the definitions of Kairos vary from narrow translations such as "particular point in time" and "specific circumstance", to wider concepts such as "situation", "occasion" and "opportunity". The narrow translation of Kairos is easily related to the rhetorical concept of Aptum, and is as such more applicable to the specific communicative situation. The wider definition however, contributes to the understanding of the ontology of rhetoric, as it clarifies the field's influence upon the

world. Not only does rhetoric construct situations with an epistemological potential, it also shares a connection with the concept of doxa (unwritten rules or joint conviction), thereby relating Kairos to practical knowledge and experience, in contradiction to knowledge in the philosophical sense. When considering the different meanings of Kairos in a Persuasive Design context, the narrow definition serves well in relation to specific design related choices, such as determining the appropriate time for initiating a persuasive strategy (i.e. triggering a specific behavior), an argument which has been raised by several researchers over the years [2] The wider definition on the other hand, supports the argument that in order to successfully select and apply a persuasive principle to the design of a technological device, the designer must beforehand acquire a fundamental understanding of the context in which the device is to be applied, and use this knowledge to create a technology which will be appropriate to the given situation.

3 Notions of Teaching and Learning

The learning perspective within the PLOT project is sought to accommodate two specific learning technologies; GLOMaker and 3ET. GLOMaker is an authoring tool for creating learning objects, whilst 3ET is an exercise tool which automatically generates grammatical exercises for language learning. Learning Objects and similar types of learning technologies, have often been criticized for not taking learning theory and pedagogy into consideration [5, 6]. They are often described as taking a "water fall approach" to learning, by which "knowledge" is presented through a technology and students are expected to learn simply from being exposed to the learning material in a different and multimodal way. One primary issue with this approach to learning is that focus is on the presentation of the learning material, rather than on the actions of the learner. In the acknowledgement that persuasive technologies by definition require interaction between the user and the technology, it appears that one of the benefits of considering persuasive design in relation to learning objects may be the promotion of a more constructivist approach to learning, in which student motivation and activity is considered a requisite for learning. By doing so, the theoretical foundation also considers the 3ET tool which by nature calls for students to engage in learning activities and build upon their existing knowledge as they practice and train new aspects of grammar.

3.1 Outcome Based Learning and the Notion of Constructive Alignment

Amongst the widely accepted approaches to constructivist learning, is the notion of Constructive Alignment (CA), which was introduced by John Biggs and Catherine Tang [7].

Biggs and Tang refer to three different levels of teaching out of which the first two are blame models, first level blaming the student and the second level blaming the teacher. The third level integrates learning and teaching and considers teaching as motivating students to use the provided learning activities in order meet the intended

learning outcome. With this third level of teaching Biggs and Tang relate their approach to teaching to the notions of outcome based learning (OBL) and constructive alignment (CA). CA is constructive in the extent that it is based on the constructivist theory that learners use their own activities to construct knowledge (or other outcomes). The alignment reflects that the learning activities and the intended outcomes must be aligned and activated in the teaching if the outcomes are to be achieved. If the intended outcome is to learn how to drive a car, focus should be placed on the activity itself, i.e. driving, rather than be focused on giving lectures on how to drive. Finally, the assessment should focus on how well the car is driven. In short, the teacher aligns the planned learning activities with the intended learning outcomes. [7]. One of the fundamental notions in CA is that the achievement of intended learning outcomes depends on a mutual responsibility between the teacher and the student. This also appears to be the primary divergence from traditional OBE, where responsibility is placed solely on the teacher. CA on the other hand acknowledges that whilst the teacher is responsible for creating the appropriate learning environment, the actual learning is something which takes place within the individual student. The teacher may inspire and guide, but in the end the student is responsible for his or her own learning. By considering learning a mutual responsibility between the teacher and the student, the notion of learning may be related to the previous argument that there must be an appropriate balance between the intentions of the designer and the intentions of the user if a persuasive design is to be successful.

4 Future Research

In spite of the novelty of Persuasive Design, it appears that the human centered perspectives described in this paper, may be a valuable contribution the field of didactics. In particular aspects of Persuasive Design may provide nuanced perspectives for teachers who aim to motivate students to actively engage in learning. Although Persuasive Design focuses on the design of interactive computer technologies, the design principles are not limited to virtual implementation. The structure of a lecture can be considered an example of the persuasive principle of tunneling, and depending on the content of the slides, teachers may include principles such as suggestion and simulation into their presentations. Finally, the rhetorical notion of Kairos which is considered one of the key concepts within Persuasive Design may also impose more nuanced reflections concerning the timing of a lecture and the facilitation of the location. Likewise the field of persuasive design is likely to be enriched by the perspectives on didactics which have been presented in this report. In particular the notion of CA may be related to the described necessity of ensuring an appropriate balance between the designer and the user of a Persuasive Technology, and CA may contribute with perspectives on how to establish this balance. The act of teaching in itself may be considered an act of persuasion. Students who attend a lecture are to some extent persuaded to change attitude towards a subject, depending on the teacher's ability to present the subject material in an appropriate manner, and to conduct the lesson in a way which upholds the engagement and interest of the students. Kairos once again

becomes a key concept as the teacher accommodates contextual changes into the planned lesson, and even the preparatory phases before the actual lecture takes place, calls for considerations concerning timing, use of location and manner in which the material is introduced.

When teaching and learning becomes computer mediated, the means to adjust and modify are altered, and must to some extent be considered to even greater detail prior to the implementation of the learning technology. System embedded adjustments dependent on student activity, or perhaps even notions of branching time, may serve as a way to ensure the persuasive characteristics of learning technologies such as those being developed through e-PLOT. Whilst the initial steps taken within the E-PLOT project have demonstrated several immediate overlaps and potential benefits of applying persuasive design strategies in the development of learning objects in particular and possibly e-learning in general, the theoretical perspectives presented in this paper also call for careful consideration regarding the contradictions between learning and persuasion, and the implications these may have not only for this specific project but also for research on Persuasive Learning on general. Terry Mayes and Sara de Freitas have argued that in order to develop good pedagogical designs, one must necessarily adopt theories of learning which go beyond the notion of CA as it is presented by Biggs and Tang [8]. Whilst Biggs and Tang do provide an applicable outline for learning, every element within the learning scenario results in implicit assumptions regarding relevance. Each intended learning outcome requires a different kind of theoretical perspective and a different pedagogical approach [8].

The word Pedagogy stems from ancient Greek, in which *Paidagogos* was the privileged slave who ensured that children made it safely to school and were not distracted on the way there. The task was simply getting the child to school at any means, and part of the privilege of Paidagogos was the authorization to corporally punish the children if they did not go freely. Naturally the notion of pedagogy has developed since, however even modern pedagogy does to some extent focus on making students do something they actually don't want to do. Not many school children long to go back to school when they are off on vacation, and not many students involved in the E-PLOT case scenarios hold an intrinsic motivation to learn 'biblical Hebrew grammar', or 'how to handle dangerous chemicals'. The students' motivation to learn is linked to a more general desire to acquire an academic degree, or remain capable of holding a specific job position. As such, pedagogy may in some ways conflict with the basic concept of persuasion, in the sense that persuasion by definition opposes manipulation, deception and force. Students may potentially be motivated by persuasive learning designs, but the process of getting the students to use the technologies may not be force-free but simply mandatory depending on how the testing and evaluation of the learning technologies is done. This taken into consideration, future research in persuasive learning must include reflections regarding the extent of persuasive elements in learning technologies and ways of ensuring the ethicality of such learning resources. This is of particular importance as design principles such as those presented in the frameworks of various researchers [9, 10], may only be persuasive if they are initiated in Kairos. A technology does not become persuasive simply by adding one or more persuasive principles, it is the ability to initiate such principles at the

right moment in time which enables them to serve a persuasive purpose. Failing to consider Kairos in the design process may not only result in insufficient persuasive designs, it may potentially also undermine the core notion of persuasive learning, as the learning technologies cannot be defined as persuasive if they are manipulative or deceptive.

Both the theoretical overlap described in this paper, and a persuasive design analysis of GLOMaker and 3ET [11] supports the often made presumption, that persuasive principles are already applied in various ways in learning. As a result, in order to fully argue that persuasive design may be an asset to the more established field of digital learning, it must be clarified how we can distinguish between persuasive learning technologies, and enriched digital learning resources. Kairos and the notion of appropriate timing may be one of the primary distinctions, as well as the aim to reinforce the balance between teacher and student intentions.

References

1. Kjær Christensen, A.-K., Hasle, P.: Classical Rhetoric and a Limit to Persuasion. In: de Kort, Y.A.W., IJsselsteijn, W.A., Midden, C., Eggen, B., Fogg, B.J. (eds.) PERSUASIVE 2007. LNCS, vol. 4744, pp. 307–310. Springer, Heidelberg (2007)
2. Glud, L.N., Jespersen, J.: Leth Conceptual analysis of Kairos for Loca-tion-based mobile devices. University of Oulu. Department of Information Processing Science. Series A, Research Papers, pp. 17–21 (2008)
3. Gram-Hansen, S.B.: Persuasive Everyware - Possibilities and Limitations. In: WMSCI. IIIS, Orlando (2010)
4. Hansen, J.B.: Den rette tale på det rette tidspunkt. Retorik Magasinet, 74 (2009)
5. Wiley, D.A.: Connecting learning objects to instructional design theory: A definition, a metaphor, and a taxonomy. In: The Instructional Use of Learning Objects (2002)
6. Dalsgaard, C.: Åbne læringsressourcer - mod en sociokulturel teori om læringsressourcer. In: Institut for Informations- og Medievidenskab, Humanistisk Fakultet. Århus Universitet, Århus (2007)
7. Biggs, J., Tang, C.: Teaching for Quality Learning at University, 3rd edn. Society for Research into Higher Education & Open University Press (2007)
8. Meyes, T., Freitas, S.d.: The role of theory. In: Beetham, H., Sharpe, R. (ed.) Rethinking Pedagogy for a Digital Age. Designing and delivering e-learning. Routledge (2008)
9. Fogg, B.: Persuasive Technology, Using Computers to change what we Think and Do. Morgan Kaufmann Publishers (2003)
10. Oinas-Kukkonen, H., Harjumaa, M.: A Systematic Framework for Designing and Evaluating Persuasive Systems. In: Oinas-Kukkonen, H., Hasle, P., Harjumaa, M., Segerståhl, K., Øhrstrøm, P. (eds.) PERSUASIVE 2008. LNCS, vol. 5033, pp. 164–176. Springer, Heidelberg (2008)
11. Gram-Hansen, S.B., Schärfe, H., Winther-Nielsen, N.: Deliverable Report D.3.2 Persuasive Learning Design Framework (2011)

Does a Hyperlink Function as an Endorsement?

Eric DiMuzio[1] and S. Shyam Sundar[1,2]

[1] Media Effects Research Laboratory, The Pennsylvania State University, USA
[2] Department of Interaction Science, Sungkyunkwan University, Seoul, Korea
ejd5050@gmail.com, sss12@psu.edu

Abstract. Websites often provide hyperlinks to other sites featuring related content. Does this imply an endorsement of the content presented on those linked sites? We answered this question with a 2 x 2 between-subjects factorial experiment in which we systematically varied source credibility (high vs. low) and source sequence (linking vs. linked). All study participants (N = 572) read a dubious story questioning the value of sunscreen, with a link at the bottom to a related story with the same claim. Results show that user perceptions of interest value (i.e., appeal) do transfer over from the linking site to the linked site. However, the source signal is quite salient when it comes to evaluations of newsworthiness and trustworthiness, with users taking note of the differences in credibility. Design implications of these findings for persuasive communications are discussed.

Keywords: Hyperlink, Link, Source, Endorsement, Persuasion.

1 Introduction

Today, a vast amount of online content is accessed using hyperlinks embedded in--and propagated across--a range of Websites that are increasingly interlinked. For example, BBC.com embeds links to related articles (published by other news outlets) within their own online articles. Users must now identify and assess multiple sources, making it harder to identify one centralized authority [4]. Theory and research suggest that users rely on a situational and/or collective/social assessment of content while they use various cues (e.g., source of an article) to invoke mental shortcuts or "heuristics" (such as "experts' statements can be trusted") [9] to evaluate online content. Certain cues on the interface, e.g., those indicating pedigree of a source or popularity of content, are known to trigger cognitive heuristics (authority and bandwagon heuristics, respectively) that bias the way readers perceive online content [2]. Given that the pedigree of a credible source serves a persuasive function, is it likely that the content suggested by this source will also be persuasive, even though it is produced by a less credible source? Is there an implicit endorsement by the linking source?

2 Literature Review

In news consumption, unless the user is highly involved with the topic, s/he will not scrutinize sources presented in a single interface [1]. Readers don't want to expend

M. Bang and E.L. Ragnemalm (Eds.): PERSUASIVE 2012, LNCS 7284, pp. 268–273, 2012.

the cognitive energy to assess concrete indicators of credibility—they will orient to the most immediate source and rely heavily on cues (e.g., authority, bandwagon) to make snap judgments regarding that source's credibility and therefore the veracity of the information accessed in that context.

Given this, it seems quite likely that the most immediate source (i.e., the source most proximate and visible) is most influential in our judgments about the credibility of the information we receive online. This is especially likely when there is a chain of hyperlinked sources involved in delivering that information, thus raising the question: are users losing sight of (or perhaps willfully ignoring) the original source as they often attribute the content (at least in part) to various sources along the source chain? Exactly how do we make source attributions to the plethora of forwarded and hyperlinked content that we consume daily?

2.1 Source Attribution

In order to evaluate a message, the receiver must attribute the message to a source. This is a particularly linear process in traditional media, which tend to feature one readily identifiable source. Online, the process is more convoluted. Any of the elements of traditional communication models (e.g. sender/presenter, medium/channel, receiver/audience) can function as a cognitively distinct source [7]. By presenting a link, is the linking source (e.g., BBC.com) supplanting the linked source? Research suggests that it is possible. Users orient towards the most immediate source in traditional media—for instance, a TV news audience's credibility assessments are primarily influenced by the news anchor's credibility [5]. The audience often ignores the fact that a team prepared the news, and orients instead to the "visible" anchorperson.

News aggregators, portals, and traditional news media company websites all function as presenters and thus as "visible" sources [7]. Each of these sources is cognitively distinct, that is to say, once a user attributes content to a source, they trigger a cascade of heuristics (machine, authority, bandwagon) regarding the source's credentials for serving as a source, in turn affecting the perceived credibility of the information provided by that source [9]. It is unclear which sources are remembered by the users and which sources exert influence on credibility and content perceptions, as users progress along a chain of sources.

Credibility assessments are increasingly based upon a social or collective assessment [4]. A decreased emphasis on traditional credibility indicators suggests that additional factors (e.g., online endorsement) may in fact play an important role in users' credibility assessments. A sequence of sources might influence the user's perception of the final source and the content it provides. By embedding hyperlinks, traditional news media companies may be lending their credibility to the target source, its contents and beyond.

3 Method

We experimentally manipulated source sequence and credibility to assess whether hyperlinking constitutes endorsement in the minds of users and thereby serves a

persuasive function. We conducted a 2 (source credibility: low vs. high) by 2 (source sequence: linking vs. linked) between-subjects, fully-crossed factorial experiment. The four conditions were fully counter-balanced in order to minimize brand and order effects. We conducted a pretest to identify sources that respondents in our sampling frame would perceive as high/low in credibility. *The New York Times* was the highest with *The Washington Post* coming in second. The *National Enquirer* was deemed least credible followed by *Star*. Webpage HTML was modified by the research team for use in the study. The content was composed by the research team and was devoid of any source cues (e.g., author bylines). All advertisements were removed and all links were disabled. Site elements such as mastheads and style sheets were preserved to maintain ecological validity.

Participants (N = 572) were recruited from a research participant pool at a large northeastern state university in the US. They received course credit for participating in the study. They were emailed a link, which randomly assigned them to a condition. They were exposed to the linking site first, with a short article titled "Vitamin D Deficiency: Is Sunscreen to Blame?" They were directed to read the content and use the link under a "From other news sites" section which exposed them to the second article titled "Stop Sunphobia: Moderate Sun Exposure Would Replenish Critically Low Vitamin-D Levels". Upon reading the stimuli, participants were immediately directed to online questionnaires where we gathered the key measures for our study.

In order to check for successful manipulation and gauge participant perception of source credibility, a 13-item, 7-point Likert scale was used [3, 6]. The following adjectives were placed adjacent to a 7-point scale, anchored at each end by (1) *describes very poorly* and (7) *describes very well*: Reliable, Biased, Credible, Unqualified, Trustworthy, Attractive, Poorly Organized, Reputable, Accurate, Interactive, Comprehensive, Valuable, and Professional.

To measure participant perception of the content provided by each source, a 22-item Likert type scale was adapted from a news-perceptions scale [8]. Items, described in the next section, were anchored between *describes very poorly* (1) and *describes very well* (7).

3.1 Index Construction

Each set of the thirteen Credibility Perception items were combined into two separate indices labeled Linking Source Credibility (Cronbach's α = .96) and Linked Source Credibility (Cronbach's α = .96).

Exploratory factor analyses were conducted on the data resulting from the 22 Content Perception items to create linking and linked versions of the following indices: "Trustworthiness" including Reliable, Believable, Comprehensive, Fair, Informative, Insightful, Objective, and Well-written (Cronbach's α = .89); "Interest Value" including Disturbing, Enjoyable, Interesting, Lively, Pleasing, and Sensationalistic (Cronbach's α = .78); and "Newsworthiness" including Clear, Coherent, Concise, Important, Relevant, and Timely (Cronbach's α = .83). The items Biased and Boring were unreliable and dropped from analyses.

In all, four indices, one relating to linked source (Credibility) and three concerned with linked content (Perceived Trustworthiness, Interest Value, and Newsworthiness of the linked story) served as the dependent variables of interest for the current study.

4 Results

The credibility manipulation was successful. For both linking, F (1, 568) = 682.41, p < .0001, and linked, F (1, 568) = 534.36, p < .0001 sources, high credibility sources were perceived to be significantly higher in credibility than low credibility sources. Linking high-credibility (M = 5.44) and linked high-credibility (M = 5.51) sources were perceived to be significantly higher in credibility than their low credibility counterparts (linking M = 3.12 , linked M = *3.39*).

Trustworthiness. General linear model analysis revealed that perceptions of linking source trustworthiness had a significant main effect, F (1, 564) = 233.41, p < .0001, on perceived trustworthiness of the linked source. Additionally, there was a significant interaction effect between linking source credibility and perceived trustworthiness of the linking content upon perceived trustworthiness of the linked story, F (1,564) = 5.09, p < .05, such that the relationship between trustworthiness of the linking (i.e., proximate) content and that of the linked (or distal) content is stronger when the proximate source is high, rather than low, in credibility. A significant three-way interaction between source sequence, source credibility, and trustworthiness, F (1, 564) = 11.26, p <.001, revealed that the relationship between perceived trustworthiness of the linking story and perceived trustworthiness of the linked story was quite strong in 3 of the 4 conditions, with the influence being weakest when the linking source had low credibility and the linked source was of high credibility. In general, user's trustworthiness perceptions of the initial content are less predictive of their perceptions of the linked content when they are exposed to sources that are discordant (H→L or L→H) in credibility.

Interest Value. Participants' perceptions of linking content's interest value had a significant main effect on their perceptions of linked content's interest value, F (1, 564) = 487.55 p < .0001. There were no interaction effects between source sequence, credibility, and interest value.

Newsworthiness. Perceived newsworthiness of the linking story had a significant main effect, F (1, 564) = 496.47, p < .0001, on perceived newsworthiness of the linked story. There was a significant 3-way interaction between linking source credibility, linked source credibility and perceived newsworthiness of the linking content upon perceived newsworthiness of the linked content, F (1, 564) = 5.57, p < .05. The relationship between newsworthiness of the linking content and that of the linked content is stronger when the linked source has low, rather than high, credibility. Perceptions of linked content's newsworthiness are more strongly dependent on the linking source's credibility when the linked source is low in credibility.

5 Discussion

At the outset, we sought to gauge the extent to which users' perceptions of a source's credibility and content are influenced by hyperlink access from another source toward which they are oriented. Our data reveal that links are indeed functional endorsements—perceptions of content (and source) were altered because users were navigating to the page using hyperlinks provided by another source. Main effects demonstrate that linked content perceptions are indeed related to linking content perceptions.

However, results also show that the positive correlation between perceptions of linking content and those of linked content is diminished for trustworthiness and newsworthiness when there is a discrepancy in the relative credibility levels of the linking and linked sources. These findings suggest that readers do in fact pay attention to sequences of sources online, and that the various layers of sourcing do not go unnoticed. When encountering discordant source credibility, users do not simply invoke the automatic rules they rely on when evaluating content. They tend to become more systematic in their processing of the linked source's message—which leads to increased scrutiny of content, resulting in evaluations that are less correlated with their evaluations of the linking source's content. If this is in fact the case, further investigation may yield other methods of instantiating a similar effect and preventing blind transfer of content perceptions. An obvious suggestion emerging from this result is that source labels ought to be prominent in online media so that users can more readily notice the differences in credibility across sources when they navigate from one site to another.

While discrepant source credibility may give the user a renewed awareness of source, the significant three-way interactions raise questions about the transfer of content perceptions. User's awareness of the layered sources did not prevent linking content perceptions of trustworthiness and newsworthiness from having a significant impact on linked perceptions. Moreover, the significant relationship between interest value of linking content and that of linked content speaks to the strong transfer effect based on content, rather than source, epitomizing the widespread phenomenon of content going viral in social media. It is clear that hyperlinks indeed function as implicit endorsements of the content they connect, even after statistically accounting for credibility differences in both linking and linked sources.

But, this could go both ways though, given that the perceptual variables are measured, rather than manipulated, in this study. Since the study administered questions for perceptions of both stories at the same time, one possibility is that perceptions of linked source and content affected those of linking source and content. Let us suppose for a moment that traditionally low-credibility publications began corroborating accounts published by high-credibility outlets and designed their site to interlink their web-presences (e.g. via a news aggregator). It is quite possible that users would begin perceiving the content on the low-credibility site as more trustworthy and newsworthy because it benefits from the halo effect (or reflected glory) of the high-credibility site to which it is linked.

In designing persuasive technologies for well-being, it is critical to ensure the veracity of the underlying health information used for recommending a particular course

of action to users. Our study suggests that users do judge this based on credibility of the communication source, with differences in relative credibility levels of different sources giving them pause and forcing a more effortful consideration of the advocated position. Yet, the interest value of the content itself seems to trump such considerations. Therefore, the general practical recommendations emerging from this study is to design interfaces such that they (1) highlight sources, ideally with credibility tags, (2) highlight the logic and mechanics of hyperlinking between content from sources of varying credibility, (3) constantly alert users to the pitfalls of heuristic cues, and (4) differentiate between content that has high interest value and that which is reliable. Designers could focus on solutions to leverage such reflected gory (e.g., pay-per-click hypertext placement) or prevent negative associations (e.g., high-credibility site embedding interstitials disclaiming association with the low-credibility site). They may also capitalize on the blind transfer of perceptions of interest value, as they devise ways to make content go viral. However, if the design goal is to prevent such a transfer, our findings suggest that showcasing the discrepancy in credibility of linking and linked sources may lead users to process the information more systematically.

Acknowledgement. This research is supported by the Korea Science and Engineering Foundation under the WCU program funded through the Ministry of Education, Science and Technology, S. Korea (Grant No. R31-2008-000-10062-0).

References

1. Kang, H., Bae, K., Zhang, S., Sundar, S.S.: Source cues in online news: Is the proximate source more powerful than distal sources? Journalism & Mass Communication Quarterly 88, 719–736 (2011)
2. Kim, H.-S., Sundar, S.S.: Using interface cues in online health community boards to change impressions and encourage user contribution. In: Proc. CHI 2011, pp. 599–608. ACM Press (2011)
3. Metzger, M.J.: Making sense of credibility on the web: Models for evaluating online information and recommendations for future research. JASIST 58(13), 2078–2091 (2007)
4. Metzger, M.J., Flanagin, A.J., Medders, R.B.: Social and heuristic approaches to credibility evaluation online. J. Communication 40, 419–439 (2010)
5. Newhagen, J., Nass, C.: Differential criteria for evaluating credibility of newspapers and TV news. Journalism Quarterly 66, 277–284 (1989)
6. Rubin, R.B., Palmgreen, P., Sypher, H.E.: Communication research measures: a sourcebook. Guilford (1994)
7. Sundar, S.S., Nass, C.: Conceptualizing sources in online news. J. of Communication 51(1), 52–72 (2001)
8. Sundar, S.S.: Exploring receivers' criteria for perception of print and online news. Journalism & Mass Communication Quarterly 76(2), 373–386 (1999)
9. Sundar, S.S.: The MAIN model: A heuristic approach to understanding technology effects on credibility. In: Metzger, M.J., Flanagin, A.J. (eds.) Digital Media, Youth, and Credibility, MIT Press (2008)

Comparative Analysis of Recognition and Competition as Features of Social Influence Using Twitter

Agnis Stibe and Harri Oinas-Kukkonen

Oulu Advanced Research on Software and Information Systems,
Department of Information Processing Science, University of Oulu
P.O. Box 3000, FIN-90014 Oulu, Finland
{Agnis.Stibe,Harri.Oinas-Kukkonen}@oulu.fi

Abstract. This paper studies how and to what extent social influence design principles can persuade people to participate in sharing their feedback. For this reason, a Twitter-based system was designed with persuasive software features of social influence at its core. The effects of recognition and competition features were tested in a pilot study in two computer rooms simultaneously. Their effects on the behavior of simulated airline travelers were compared. The main result of this study provides evidence of several positive effects, especially regarding the persuasive powers of recognition in a system's design.

Keywords: Social influence, persuasive software features, Twitter, recognition, competition, user participation, behavior change.

1 Introduction

Mobile devices and pervasive technologies are integrated with physical space in order to facilitate social interaction [1], and often these environments are complemented with interactive interfaces for motivating people to change their behavior [2]. Socially active users, in combination with improved public environments, provide various opportunities for businesses to allow their customers to participate by providing feedback. For example, airline travelers waiting for their flights can be approached with questions about their satisfaction with the services provided in an airport or about their wishes and concerns related to traveling. This can make such travelers feel more engaged.

Theories from the social sciences, such as social cognitive theory (SCT) [3], describe fundamental human capabilities that define who people are and emphasize that they learn from observing others who are performing the behavior. In this case, airline travelers may learn by observing others providing feedback in an airport and then may change their own behavior accordingly by getting more involved in sharing their thoughts. Such feedback collection can be automated by an information system, but its persuasiveness can be enhanced by the design principles of the Persuasive Systems Design (PSD) model [4]. This paper describes a pilot study that, using such a system, explores and compares the effects of recognition and competition as persuasive software features on user participation in sharing feedback.

M. Bang and E.L. Ragnemalm (Eds.): PERSUASIVE 2012, LNCS 7284, pp. 274–279, 2012.

2 Research Framework

To study the area of social interaction, the previously conducted research suggests looking at SCT, which "analyses social diffusion of new styles of behavior in terms of the psychological factors governing their acquisition and adoption and the social networks through which they spread and are supported" [5]. Particularly, SCT describes the human capability for self-regulation [6], which includes self-monitoring behavior, making judgments in relation to personal standards, and experiencing affective self-reaction.

The framework for PSD [4] defines key issues behind persuasive systems, the process model, and the design of system features. It contains a model constructed as a meta-level model that requires being used together with related theories [7]. In this case, such theories focus on a historically long-standing problem of social psychology: how and why the behavior of one person affects the behavior of another. In this study, the following two features of the social influence category of the PSD model [4] are examined and compared: *recognition*, the design principle aimed to increase the likelihood of a person or group adopting a target behavior by offering public recognition of an individual or group through a system, and *competition*, which characterizes the potential of a system to motivate users to adopt a target attitude or behavior by leveraging humans' natural drive to compete.

The related research discussed previously provides support for the integration of the social cognitive perspective [3, 5, 6] and the persuasive software features of social influence [4] to study user participation in sharing feedback. For that reason, a research framework for the study is proposed in Fig. 1.

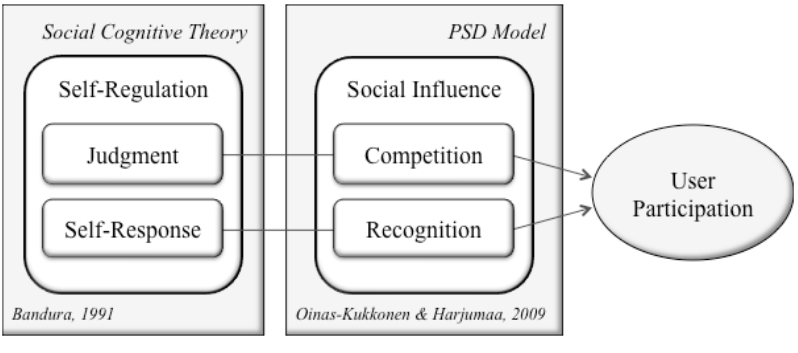


Fig. 1. The research framework, representing related theories from the social sciences (*left column*), connected constructs from the social influence category of the PSD model (*middle column*), and the study domain of behavior change (*oval to the right*)

The research framework was built upon the self-regulation construct of SCT, wherein the judgment sub-construct conveys ideas related to competition and wherein the self-response sub-construct supports principles of recognition [6]. Therefore, competition and recognition features of social influence were connected with lines to their corresponding sub-constructs of self-regulation on the left, but arrows on the right were added to represent their influence on user participation. See Fig. 1.

3 System Description

For the purposes of this study, a system containing recognition and competition as persuasive software features was created to alter behavioral patterns related to user participation in sharing feedback. The system was developed on top of the Twitter (twitter.com) platform, which previously has been verified as being effective for user engagement [8]. Twitter is a microblogging web service with increasing popularity in recent years, which allows registered users to post messages that are limited to a hundred and forty characters.

The designed system had both a public interface and an interface for an administrator. Questions were added to the system through the administrator's interface, whereupon a special hashtag (#) was assigned to each entry. Based on these hashtags, the system was able to select all relevant messages ("tweets") from the actual Twitter feed in order to store them in a database for further processing. The collected data was frequently aggregated by the system and then displayed on the public interface. The loop principle determined a continuous rotation of questions and related data on the public interface at an admin-specified frequency. The layout of the public interface included a question, the related hashtag, all tweets containing this hashtag, the number of active participants, the sum of collected tweets, and either a list of top responders (the competition feature of the study) containing most active users or a space for publicly recognizing well-performing users (the recognition feature of the study) depending on users' individual achievements in tweeting.

4 Research Setting

A pilot study employing the system was conducted in a class setting where users were asked to pretend they were airline travelers waiting for their departures in an airport. The output of the system was projected on a big screen in front of the class, and the users responded via Twitter from desktop computers or mobile devices. The study was conducted for 30 minutes. Six questions related to airline-travel issues were added into the system incrementally. At the beginning of the study, two questions rotated in a loop on the big screen. After ten minutes, another two questions were added, and after another ten minutes, the last two questions were added. The displayed information was automatically refreshed every fifteen seconds.

All participants of the pilot study were computer science students in a graduate program who were enrolled in a course about information and communication technologies and behavior change. They were randomly divided into two groups and were placed in two separate computer rooms. One computer room projected the public interface of the system with the recognition feature and the other computer room showed the public interface of the system with the competition feature.

Immediately after the pilot study, all users were required to fill in an online questionnaire containing mandatory questions, which were mainly aimed at eliciting quantitative data about experiences using the system. It employed a seven-point Likert scale for assessing attitudes, the first option being "strongly disagree," the last option being "strongly agree," and the middle option being "undecided."

5 Data Analysis and Results

Twenty-four males and thirteen females participated in the pilot study (N=37). Most of them (n=32, 87 percent) were between 20 and 29 years of age. Nearly all respondents (n=34, 92 percent) agreed that the system was useful for feedback collection, and a large number of the respondents (n=31, 84 percent) thought that the system effectively encouraged users to participate.

The stem-and-leaf method was used to verify the normality of distribution throughout the data set, and a further analysis was carried out using an independent samples t-test provided by SPSS (Statistical Package for the Social Sciences) statistical software.

5.1 Comparing Recognition and Competition as Persuasive Software Features

The performed t-test analysis comparing users who saw the recognition feature (n=18) and those who saw the competition feature (n=19) revealed significant differences in answers to several questions. See Table 1 for key findings.

Table 1. T-test results for the effect of the displayed persuasive software feature (*recognition vs. competition*) on the dependent variables *(** p<0.01; * p<0.05)*

Item	Recognition	Competition	t-value	df	p
Twitter is a powerful tool to call for action outside the virtual world.	5.50	4.32	2.937	35	.006**
I believe that the system would work well in a real airport.	5.56	4.47	2.775	35	.009**
I think that the system is effective for encouraging users to participate.	6.11	5.11	2.570	35	.015*

Users who had seen the recognition feature were significantly more affirmative regarding the power of Twitter to call for action outside the virtual world ($t(35)=2.937$, $p=.006$), they had higher levels of belief that the system would work well in a real airport ($t(35)=2.772$, $p=.009$), and they were more positive about the system's effectiveness in terms of encouraging users to participate ($t(35)=2.570$, $p=.015$). These findings provide support for the increased persuasive capacity that the recognition feature had in this type of setting.

5.2 Comparing Users Based on Their Appearance on the Big Screen

Users who had seen themselves on the big screen being recognized with special titles or who had seen themselves in the list of top responders were significantly more affirmative regarding the ability to monitor their performance ($t(33)=4.512$, $p=.000$) and their motivation to produce more tweets ($t(33)=2.352$, $p=.025$). See Table 2.

Table 2. T-test results for the effect of the dimension of whether users had seen themselves recognized with special titles or seen themselves in the list of top responders (*yes vs. no*) on the dependent variables *(** p<0.01; * p<0.05)*

Item		Yes	No	t-value	df	p
Displaying public recognition or the top responders helped me to monitor my performance.	*All*	5.44	3.25	4.512	33	.000**
	Recognition	5.54	3.50	3.427	15	.004**
	Competition	5.36	3.00	2.977	16	.009**
Tweets provided by others on the big display encouraged me to come up with my tweets.	*All*	Non-significant difference				
	Recognition	5.69	5.00	3.323	12	.006**
	Competition	Non-significant difference				
Displaying public recognition or the top responders motivated me to produce more tweets.	*All*	5.00	3.75	2.352	33	.025*
	Recognition	5.38	3.50	2.409	15	.029*
	Competition	Non-significant difference				

More detailed analysis revealed that users who had seen themselves being recognized with special titles were not only more affirmative regarding the ability to monitor their performance ($t(15)=3.427$, $p=.004$), and a motivation to produce more tweets ($t(15)=2.409$, $p=.029$), but they were also more encouraged to come up with their own tweets based on the tweets of others ($t(12)=3.323$, $p=.006$). On the contrary, those users who had seen themselves in the list of top responders were more affirmative only in regards to the ability to monitor their performance ($t(16)=2.977$, $p=.009$). These findings complement the idea that the recognition feature had more persuasive powers comparing to the competition feature in this setting.

6 Discussion

The results from the pilot study provided evidence regarding several positive effects of social influence design principles on user participation. Almost all respondents considered the system to be useful for feedback collection, and very many thought that the system was effective for encouraging users to participate.

The comparative analysis of recognition and competition as persuasive software features demonstrated support for the view that the recognition feature is especially persuasive in influencing users' beliefs about the powerfulness of Twitter to call for action outside the virtual world, about the effectiveness of the system to involve users to participate, and about the system's expected performance in a real airport. The analysis of users who had seen themselves being recognized with special titles compared to users who had seen themselves in the list of top responders revealed that the recognition feature was more compelling than the competition feature in terms of helping users monitor their performance, motivating them to be more active in tweeting, and encouraging them to come up with their own tweets.

In summary, the results of the study demonstrated positive effects of two social influence features on user behavior, particularly emphasizing that the recognition feature had more persuasive powers than the competition feature within the setting.

7 Conclusions

The study explored and compared the effects of recognition and competition as persuasive software features on user participation in sharing feedback. The research framework based on related research was proposed, and a system with both features at the core of the implementation was developed. The results of the pilot study demonstrated several positive effects of the social influence features on user behavior, especially emphasizing the persuasive capacity of the recognition feature.

The study has several limitations. The experiment was conducted in a class setting where users were able to watch others performing the behavior. The participants represented a somewhat limited sample in terms of age and education, which makes the generalizability of the findings questionable. The pilot stage did not include a control group that used the system without any of the two features.

Further research should focus on field testing the features. The system can be tested in other kinds of transportation hubs or public environments as well. In the future, the system can also be complemented with other persuasive software features from the social influence category of the PSD model [4], and they can be tested in parallel to enable a comparison of their effects, either separately or in combinations.

Acknowledgments. We would like to acknowledge our colleagues Payam Hossaini, for developing the system, and Pasi Karppinen, for assistance in conducting the pilot study. The study was supported by the Foundation of Nokia Corporation, as well as by the Someletti research project on Social Media in Public Space, grant 1362/31, and the SalWe Research Program for Mind and Body, grant 1104/10, both provided by Tekes, the Finnish Funding Agency for Technology and Innovation.

References

1. Eriksson, E., Hansen, T.R., Lykke-Olesen, A.: Reclaiming Public Space: Designing for Public Interaction with Private Devices. In: Proceedings of the 1st International Conference on Tangible and Embedded Interaction, pp. 31–38. ACM Press, New York (2007)
2. Mathew, A.P.: Using the Environment as an Interactive Interface to Motivate Positive Behavior Change in a Subway Station. In: CHI Extended Abstracts on Human Factors in Computing Systems, pp. 1637–1640. ACM Press, New York (2005)
3. Bandura, A.: Social Foundations of Thought and Action: A Social Cognitive Theory. Prentice-Hall, Englewood Cliffs (1986)
4. Oinas-Kukkonen, H., Harjumaa, M.: Persuasive Systems Design: Key Issues, Process Model, and System Features. Commun. Assoc. Inform. Syst. 24(28), 485–500 (2009)
5. Bandura, A.: Social Cognitive Theory of Mass Communication. Media Psychology 3(3), 265–299 (2001)
6. Bandura, A.: Social Cognitive Theory of Self-Regulation. Organizational Behavior and Human Decision Processes 50, 248–287 (1991)
7. Oinas-Kukkonen, H.: Requirements for Measuring the Success of Persuasive Technology Applications. In: Proceedings of the 7th International Conference on Methods and Techniques in Behavioral Research, Article 11, 4 pages. ACM Press, New York (2010)
8. Junco, R., Heiberger, G., Loken, E.: The Effect of Twitter on College Student Engagement and Grades. J. Comp. Assisted Learning 27(2), 119–132 (2011)

Author Index