Shlomo Berkovsky
Jill Freyne (Eds.)

LNCS 7822

Persuasive Technology

8th International Conference, PERSUASIVE 2013
Sydney, NSW, Australia, April 2013
Proceedings

Lecture Notes in Computer Science 7822

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

Shlomo Berkovsky Jill Freyne (Eds.)

Persuasive Technology

8th International Conference, PERSUASIVE 2013
Sydney, NSW, Australia, April 3-5, 2013
Proceedings

Springer

Volume Editors

Shlomo Berkovsky
NICTA, Alexandria, NSW 1435, Australia
E-mail: shlomo.berkovsky@nicta.com.au

Jill Freyne
CSIRO, Epping, NSW 1710, Australia
E-mail: jill.freyne@csiro.au

ISSN 0302-9743 e-ISSN 1611-3349
ISBN 978-3-642-37156-1 e-ISBN 978-3-642-37157-8
DOI 10.1007/978-3-642-37157-8
Springer Heidelberg Dordrecht London New York

Library of Congress Control Number: 2013933234

CR Subject Classification (1998): H.5.1-3, H.4.1-3, H.3.4-5, H.4-5, J.3, I.2.10-11, K.4.2-3

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

This volume comprises the proceedings of the 8th International Conference on Persuasive Technology (Persuasive-2013), which took place in Sydney during April 3–5, 2013. The Persuasive Technology conference series is the primary meeting for researchers and practitioners interested in how interactive technologies and services can influence people's attitudes and support positive behavior change. The conference has taken place annually since 2005 and provides a forum for the exchange of ideas and the presentation of research conducted worldwide.

Persuasive-2013 was chaired by Shlomo Berkovsky of the National ICT Australia (NICTA, www.nicta.com.au) and Jill Freyne from the Commonwealth Scientific and Industrial Organisation (CSIRO, www.csiro.au). They were assisted by an international Program Committee of 29 leading figures in the persuasive research community. The conference solicited long papers of up to 12 pages in length, which represent original reports of substantive new research. In addition, the conference solicited short papers of up to six pages in length, whose merit was assessed in terms of originality and importance.

Persuasive-2013 received 47 submissions from all over the world, with large numbers from Europe, the United States, and many Australian submissions. Each submission was reviewed by three or more reviewers and after a rigorous review process, 16 long papers and 12 short papers were selected for publication. This resulted in an acceptance rate of 34% for long papers and 25.5% for short papers.

The program of Persuasive-2013 covered a wide range of topics, with papers covering not only typical persuasive domains like health and environment, but also work covering emerging research topics, such as data safety and evaluation of persuasive technologies. The conference program also included three workshops that were selected by Jaap Ham from the Eindhoven University of Technology. The following workshops were organized: the International Workshop on Behavior Change Support Systems (BCSS), the Workshop on Personal Data: Designing for the New Currency, and the Workshop on Human Behavior for Industrial Safety (HBIS). Persuasive-2013 also included a doctoral consortium—a forum for PhD students to get feedback and advice from leading researchers in the persuasive community.

Two high-profile keynote speakers were invited to share their influential contributions in the field of persuasive technology. The conference program included keynote talks by Anthony Jameson from the German Research Center for Artificial Intelligence (DFKI) and Anind Dey from the Carnegie Mellon University. We are most grateful to both speakers for taking time out of their busy schedules to come to Sydney and participate in Persuasive-2013.

It would not be possible to hold the conference without the Program Committee and the sponsors. We deeply acknowledge the conscientious work of the

Program Committee members and the additional reviewers, who helped us to assure the quality of the accepted papers. We also gratefully thank our sponsors, NICTA and CSIRO, who helped us with funding and organizational expertise. We appreciate the efforts of the Publicity Chair, Sarvnaz Karimi from CSIRO. Finally, we want to acknowledge the use of EasyChair that simplified the review process and the preparation of the proceedings.

We hope that you find the papers in these proceedings interesting and stimulating.

January 2013

Shlomo Berkovsky
Jill Freyne

Organization

Program Committee

Magnus Bang	Linköping University, Sweden
Shlomo Berkovsky	NICTA, Australia
Robert Biddle	Carleton University, Canada
Samir Chatterjee	Claremont Graduate University, USA
Nadja De Carolis	Università di Bari, Italy
Peter De Vries	University of Twente, The Netherlands
Rosta Farzan	Carnegie Mellon University, USA
Alexander Felfernig	Graz University of Technology, Austria
Susan Ferebee	University of Phoenix, USA
Bj Fogg	Stanford University, USA
Jill Freyne	CSIRO, Australia
Floriana Grasso	University of Liverpool, UK
Ulrike Gretzel	University of Wollongong, Australia
Rosanna Guadagno	University of Alabama, USA
Anton Gustafsson	IT-University of Goteborg, Sweden
Jaap Ham	Eindhoven University of Technology, The Netherlands
Curtis Haugtvedt	Ohio State University, USA
Stephen Intille	Northeastern University / MIT, USA
M Sriram Iyengar	University of Texas Health Science Center at Houston, USA
Maurits Kaptein	Eindhoven University of Technology, The Netherlands
Tom MacTavish	Illinois Institute of Technology, USA
Martha G. Russell	Stanford, USA
Paul McKevitt	University of Ulster, UK
Cees Midden	Eindhoven University of Technology, The Netherlands
Harri Oinas-Kukkonen	University of Oulu, Finland
Eva L. Ragnemalm	Linköping University, Sweden
Anna Spagnolli	University of Padova, Italy
Oliviero Stock	FBK-irst, Italy
Nava Tintarev	University of Aberdeen, UK
Julita Vassileva	University of Saskatchewan, Canada
Fahri Yetim	University of Siegen, Germany
Johan Åberg	Linköping University, Sweden

Table of Contents

Persuasive Technology or Explorative Technology?

Anind Dey

Carnegie Mellon University
6323 Walnut st, Pittsburgh, 15206, USA
anind@cs.cmu.edu

Abstract. Persuasive technology is technology that is intended to change attitudes and/or behaviors using persuasion. An issue with such technology is it pre-supposes that individuals must be convinced to change their attitude or behavior. In contrast, explorative technology is technology that is intended to help users explore and understand information about themselves to support self-reflection and to identify opportunities for behavior change. In this talk, I will describe the differences between such technologies, and argue for the need to invest more effort in explorative technologies through the discussion of a series of case studies.

S. Berkovsky and J. Freyne (Eds.): PERSUASIVE 2013, LNCS 7822, p. 1, 2013.

How Can Persuasive Technology Help People Choose for Themselves?

Anthony Jameson

German Research Center for Artificial Intelligence (DFKI)
Saarbrücken, Germany
jameson@dfki.de

Abstract. Persuasive technology is used when it is known in advance what the person in question should be persuaded to do. The job of helping people to choose for themselves what they want to do, in accordance with their own values, tastes, past experience, and capabilities, has been left to other types of interactive system, such as those for decision support or recommendation. But there are reasons why persuasive technology researchers might consider applying their skills to the challenge of helping users choose for themselves:

1. A lot of the innovative techniques developed in persuasive technology can be adapted to yield new ways of supporting choosing. Examples are techniques for monitoring behavior, for simulating the consequences of possible actions, for enforcing commitment strategies, for mediating social influence, and for communicating with users at opportune times and places.
2. While users of persuasive technology are doing what they have been persuaded to do, they often need to make nontrivial personal choices about exactly how to do it; their overall success and satisfaction will be affected by how well they make these choices.

This talk argues for these claims with reference to concepts and results from the psychology of everyday decision making, and it illustrates them with examples from past and ongoing research and practice.

S. Berkovsky and J. Freyne (Eds.): PERSUASIVE 2013, LNCS 7822, p. 2, 2013.

Software Architecture Design for Health BCSS: Case Onnikka

Tuomas Alahäivälä[1,2], Harri Oinas-Kukkonen[1], and Terhi Jokelainen[2]

[1] University of Oulu, Department of Information Processing Science
P.O. BOX 3000 FI-90014 Oulu, Finland
[2] University of Oulu, Clinical Research Center, Department of Internal Medicine
P.O. BOX 5000 FI-90014 Oulu, Finland
{tuomas.alahaivala,harri.oinas-kukkonen,
terhi.jokelainen}@oulu.fi

Abstract. Behavior change support systems (BCSS) are a specific type of persuasive systems. They demand longer time spans and a very tight coupling with individual users' multiple real-life contexts. However, in most cases research into these systems has described technical artifacts at such a general level that important implementation details such as the software architecture have been ignored. In this paper, we will present a software architecture design for a full-fledged BCSS. The architectural style suggested defines a layered architecture and its key system components. The architecture has been implemented in a real-life BCSS for supporting weight loss and maintenance in order to prevent health problems such as metabolic syndrome. The system development process and the selection of implemented persuasive features was carried out by utilizing the persuasive systems design model. The lessons learned and the architecture presented in this paper can be used in further software engineering research regarding persuasive and behavior change support systems.

1 Introduction

Recently, behavior change support systems (BCSSs) have been introduced as objects of study within the persuasive technology field [1]. These systems are designed to form, alter, or reinforce attitudes, behaviors, or acts of complying without using deception or coercion [2]. The interests in BCSS research include the approaches, methodologies, processes, and tools for developing such systems, as well as studying their possible impacts. One of the most prominent domain areas for BCSS applications is promoting behavior change for improved health and healthier lifestyles [2].

Obesity and its co-morbidities represent one of the major public health problems globally. To prevent health problems caused by overweight, such as metabolic syndrome, a sustainable change in an individual's health behavior is required. The effectiveness of nutritional and weight counseling varies: highly effective personal counseling is too expensive for wide use, while distributing information might be cost efficient but not effective enough for sustainable weight loss. Therefore, new cost-efficient counseling methods are needed. Because of this, web-based weight loss and

S. Berkovsky and J. Freyne (Eds.): PERSUASIVE 2013, LNCS 7822, pp. 3–14, 2013.

maintenance software systems have become an important research area in health behavior change. These systems may utilize many different kinds of persuasive features [3-5]. However, there is only a little research on how they should be implemented.

Although BCSSs have emerged as an important research area, the descriptions of the systems in most cases have been presented in too general level [3], [5]. When describing a BCSS, the persuasion context, containing many design issues related to the chosen technology, its use, and the users, should be covered comprehensively [2]. With black-box thinking of software systems without an actual description of the implementation it would be difficult to argue any generalizable results. Through conceptual designs such as software architectures, describing abstractions of the components within a system and their relationships, BCSS and other persuasive systems research can be advanced from proof-of-concepts to actual software development guidelines. There have not been many prior studies on persuasive system architectures, but those that have been done include an architecture design for in-home monitoring technology to send persuasive messages to elderly adult diabetics [6] and a framework for designing intelligent healthcare self-management systems [7]. On the other hand, we are more interested in developing generalizable software architecture for BCSSs to be utilized independently of the problem domain. This paper describes a proposal for such architecture and its implementation in a weight loss and maintenance intervention system.

Section 2 will discuss central concepts related to software architectures and persuasive systems design. Section 3 will introduce key architectural components for modeling the contexts for persuasion. Section 4 will describe analytical and implementation considerations of these in the case of BCSS for weight loss and maintenance.

2 Background

2.1 Persuasive Systems Design

In their persuasive systems design (PSD) model, Oinas-Kukkonen and Harjumaa [8] stated that the development of persuasive systems requires three steps: understanding the key design issues related to persuasive systems, analyzing the persuasion context, and designing the system qualities. In the PSD model, the fundamental issues concerning persuasive systems based on prior socio-psychological theories have been summarized in *seven postulates*: information technology is never neutral, people like their views about the world to be organized and consistent, direct and indirect routes are key persuasion strategies, persuasion is often incremental, persuasion through persuasive systems should always be open, persuasive systems should aim at unobtrusiveness, and persuasive systems should aim at being both useful and easy to use.

The *persuasion context* analysis includes recognizing the intent of persuasion, the persuasion event, and the strategies in use [8]. Acknowledging the intent includes determining who is the actual persuader. Since computers don't have any intentions of their own, the source of persuasion in a system is always one of those who create, distribute, or adopt the persuasive technology [9]. Analyzing the intent also covers defining the change type [8]. The outcome/change design matrix [2] defines the three

potential, successful voluntary outcomes of BCSSs as formation, alteration, or reinforcement of attitudes, behaviors, or compliance. As for understanding the persuasion event, the contexts of use, the user, and the technology should be recognized [8]. The use context covers the characteristics of the problem domain in question, the user context includes the differences between the individuals, and the technology context contains the technical specifications of a system. Finally, identifying the persuasion strategies includes analyzing the persuasive message that is attempted to convey and the route, whether direct or indirect, that is used to reach the persuadee. It is important to realize that the persuasion context can be fleshed out as software architecture.

Concerning the design of the actual *software features* of persuasive systems, Oinas-Kukkonen and Harjumaa [8] have proposed four categories of design principles: primary task, dialogue, system credibility, and social support. These design principles may function as guidelines for determining software requirements, as well as an evaluation method for persuasive systems.

2.2 Software Architectures

According to a definition by Fielding [10], software architecture is an abstraction of runtime elements of a software system during some phase of its operation. Complex systems often contain many different levels of abstraction and different phases of operation, each with their own architectures. Fundamentally, a software architecture describes a configuration of architectural elements and their interactions with each other.

The architectural elements in software systems can be distributed into components, connectors, and data [11]. Components are abstract units consisting of software instructions and an internal state, supplying the transformation on the data via an interface. At the architectural level, a component is defined by the functionalities it provides in its interface rather than its internal software implementations. Connectors are elements that enable communication between components by transferring data between the interfaces. The data elements contain the information that is transferred among different components, be it messages, parameters, or serialized objects. The structure of the relationships among these architectural elements during a period of system runtime is labeled as configuration [11].

Since architectures describe all the functional and nonfunctional properties of a system, it could be challenging to compare architectures for different styles of systems and environments directly. Therefore, it is usually desirable to pursue a design for a more generalizable architectural style to more easily categorize different architectures and define their mutual characteristics. An architectural style is a coordinated set of constraints, restricting the roles of different software elements and the allowed relationships among them. Styles describe the abstraction of the relevant interaction between the components, and ignore the implementation details of the architecture. Novel architectures can then be constructed as instances of different architectural styles. Architecture may as well be composed of different styles in its different aspects, or hybrid styles can be composed from the prior designs [10].

3 Architecture for Modeling the Persuasion Context

As a basis for outlining the architecture for BCSSs and the distribution of a system's architectural components, we adopt the persuasion context presented in the PSD model [8]. When first the desired intent, then the event, and finally the strategies are carefully identified and analyzed, the implementation of the system can take place through use context, user context, technology context, message, and route in order to fulfill the persuasion intent. These concepts can be actualized in the architectural components of use context, user context, user-system interaction, social interaction, and system-mediated messages (see Fig. 1). The described architectural layer separates between the conceptualization and the technical backend and supports the implementation of the desired persuasive features in different programming environments.

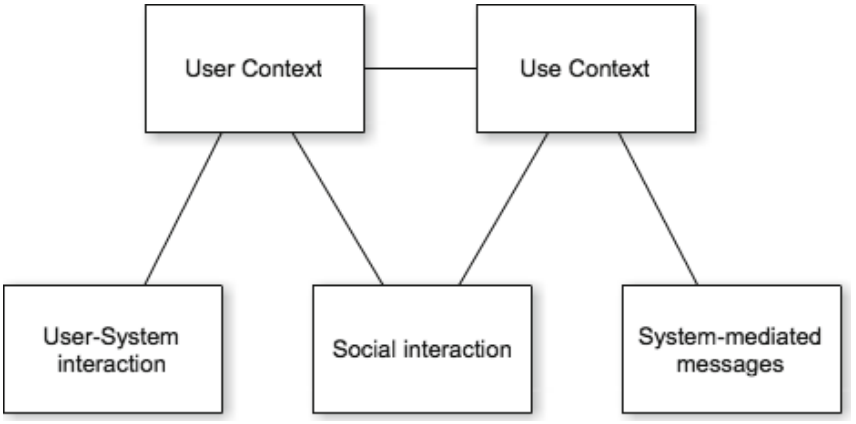


Fig. 1. Architectural components of persuasion context

The analysis of the persuasion context should begin with the analysis of the persuasion intent, including the recognition of the persuader and the desired change type [8]. The desired outcome and the change type can be derived from the designated outcome/change design matrix [2]. In the next analysis, the persuasion event contains the system's user, use, and technology contexts [8]. Since the matters related to technological context are dependent on each implementation, but on the other hand, are provided by the surrounding system itself, the software features related to the persuasion event can be implemented in the corresponding software components of user context and use context. To accomplish a more explicit separation of different software functionalities, the specific features in the use context related to social interaction and user-system interaction can be actualized as their own distinct architectural components. Persuasion strategies, actualized by the persuasive messages and their routes, can be implemented as the system-mediated messages component, to be presented to the appropriate users in their appropriate contexts of use. More precisely:

1. The software features of the *user context* component should contain all the relevant information of the user, based on which the appropriate persuasion strategies can be put into action. A wide variety of differences between individuals' interests, needs, goals, motivations, abilities, attitudes, commitment, consistency, compromises, lifestyles, persistence of change, cultural factors, deep-seated attitudes, social anchors, and personality can be addressed in the modeling of the user context software element, giving the ability to present personalized and/or tailored content for most efficient persuasion [8].
2. The features of the *use context* component enable the environment, including the graphical user interface and the back-end data storage, to convey the appropriate persuasive strategies to the users. The implementation of this component should adapt to the problem domain in question [8].
3. Setting goals and finding means to accomplish them in a systematic and effective way should be actualized in the *user-system interaction* components, allowing the system to record data about its users and the users to submit data into the system as well as to monitor their progress towards the desired change based on this.
4. An important aspect in modern web services, interaction between users for giving and receiving peer support, should be implemented in the software component for *social interaction*. Its software features should allow discussion to be raised among the desired user group.
5. The software component for the *system-mediated messages* define the content to be provided for the users in the selected form that is intended to result in behavior change. Depending on the situation, either direct or indirect routes may be used [8]. The provided messages are then presented in the appropriate use context.

This proposed architecture can be used to implement the conceptualization of a BCSS as lower-level software components, independently of the problem domain or programming environment. Utilizing the architecture ensures that the use, user, and technology contexts, message, and route are actualized in the resulting system, provided that the desired persuasion intent has been first recognized. For most efficient persuasion, the appropriate design principles should be utilized in the implementation of the individual software components.

4 Case Onnikka

The suggested architecture for persuasion context was actualized in the development of the Onnikka weight loss and maintenance intervention system as a part of studying the prevention of metabolic syndrome through new lifestyle intervention methods (see Fig. 2). The name refers to a local dialectal word meaning “a bus,” creating a metaphor of the behavior change process as a journey across waypoints. The main features in Onnikka are a channel to receive weekly content provided by health professionals, and the possibility to set goals and follow the individual’s progress within the system by submitting personal entries related to self-monitoring. The planned length of 52

weeks long intervention program is provided in order to intensively support individual's weight loss process and persevering adoption of healthy lifestyle. The development followed the persuasive systems design [8] process.

Fig. 2. The front page of the Onnikka BCSS

The individual software components were modeled according to the proposed architectural design. The key components are represented in Fig. 3.

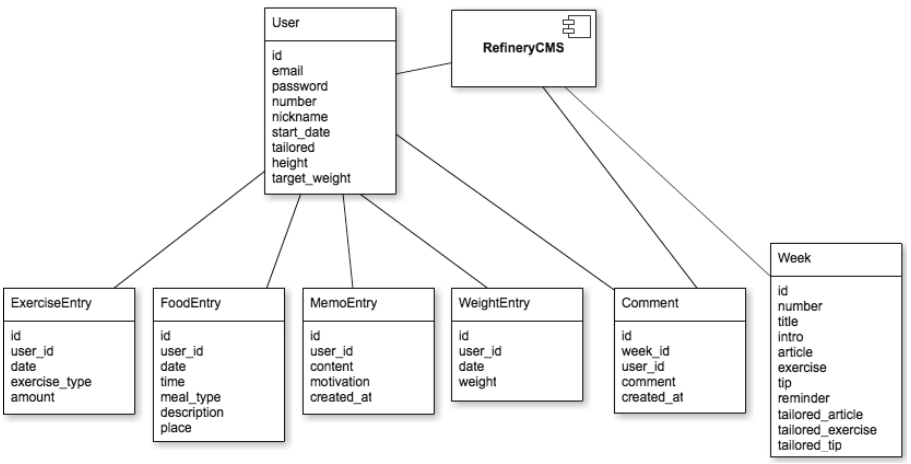


Fig. 3. System resources in Onnikka

4.1 Analyzing the Persuasion Context

4.1.1 User Context

The potential users of Onnikka are those struggling with overweight, which exposes them to diseases such as metabolic syndrome. Although pursuing a similar behavior change, the user group is not homogeneous; rather, it includes individuals with different backgrounds, goals, and motivations. The differences between individuals, primarily the starting time of the intervention and the possible need for additional eating behavior counseling, were considered when designing the software components related to the user context. One of the most important aspects of the user context was minding each user's phase of intervention, so that they could be provided the proper information at the right time. The users sharing the same starting date would belong to the same group of users during the whole intervention, being able to interact with each other. In addition, to setting conscious performance goals, the users are encouraged to set their own target weight.

It has been claimed that a successful health behavior change at the individual level requires receiving targeted and tailored health information [12]. In addition, enhancing an individual's eating behavior is one of the major factors in maintaining weight loss [13]. Since not all persons exposed to the metabolic syndrome have the specific need for counseling in eating behavior, the informational content and the software features related to enhancing eating behavior were designed to be only accessible to a certain group of tailoring-enabled users based on their behavioral profiles. This group would receive additional health information and exercises on enhancing their eating habits.

4.1.2 Use Context

The use context, containing the problem domain dependent design issues (here weight loss and management), is manifested as the different aspects of the web-based user interface of the system (see Fig. 2). We will next describe a scenario that takes into account the most essential features of use context, including the home page, the weekly information page, and the self-monitoring entries.

The home page, shown in Fig. 2, provides a personalized starting point for the user to access the different functionalities of the system. The user is reminded and suggested to submit her self-monitoring entries, as well as to read the weekly contents. She is notified about new comments in the discussions, and user statistics such as the overall weight loss by all users in the same group are shown as cues for social support. The home page also shows a progress bar visualizing the remaining time in the intervention with supportive waypoints..

In weekly contents, the user is presented with health information and exercises according to her phase of intervention and tailoring preferences. Each week contains a weekly article, exercise and a short tip. The weeks also contain a comment space for discussion among users. The user can browse back over the past weeks to re-read earlier contents. There is also an index page for the available weeks, sorting them by their subjects. Healthcare professionals provide the articles, exercises, and tips used.

The self-monitoring section is also a major functionality of the system. For tracking her change progress, the user can submit consistent entries about her weight loss

progress, memo notes about her feelings during the intervention, and her eating and exercise habits. It is suggested that the user weigh herself once a week on the same weekday and submit the result. Persuasive messages are conveyed in the submission process. The submitted entries can be browsed in table form or as a visualized graph. It is also suggested that the user write "memos," short messages related to her current thoughts and feelings about the intervention, and that she monitors her weekly exercise amounts. She may then submit entries on her exercise efforts, describing the type of exercise, the level of strain, and the amount of exercise done. The overall exercise performed during the ongoing week is presented for comparison with the recommendations from health professionals. The user is asked to keep a food diary, a tool often used in weight counseling. The submitted food entries contain information about the meal type, eating time, description, and the eating-place of a meal. After adding a meal, the user can tag an individual meal as "good" or "unnecessary," resulting in reflection on her eating habits.

4.1.3 Technology Context

The technological implementation of Onnikka was carried out using modern web technologies in order to produce a stable and highly responsive system that users could access on their desktop or mobile devices. In the development of Onnikka, we utilized two architectural styles commonly employed in web services development and extensively supported by modern web frameworks: model/view/controller and representational state transfer. The final implementation was carried out using the Ruby on Rails web framework.

The *model/view/controller (MVC)* design pattern is a widely used methodology to distribute the software functionalities into separate interactive application components: models, views, and controllers [14]. Models are the domain-specific simulations or implementations of the application's central structure; views display graphical representations of the data requested from the models; and controllers contain the interface between the associated models and views, as well as handling the user interaction from the input devices. Using the MVC design pattern makes developing and maintaining an application easier, since the application's user interface can easily be modified regardless of the data structures and the business logic within. However, Leff and Rayfield [15] have stated that since the MVC pattern is expressed in terms of an interactive application running in a single address space, applying it to web applications can be challenging since they are often deployed in various client/server architectures. They present the concept of flexible web-application partitioning to enable partition-independent web applications to be built. A possible deployment architecture supporting such partitioning is the thin-client approach, in which the model and the controllers reside in a single address space on the server, and generate the views to be rendered on the client side.

Representational state transfer (REST) is a "coordinated set of architectural constraints that attempts to minimize latency and network communication, while at the same time maximizing the independence and scalability of component implementations" [10]. It has become the de facto standard for service design in Web 2.0 applications, although REST is not specifically a defined W3C standard, but rather a widely

used design technique with several interpretations. In "RESTful" services, the application's resources are identified and resolved by a particular uniform resource locator (URL), describing the target protocol, host address, application path, resource type, and resource identifier. As such, the application's data content can be accessed by performing particular hypertext transfer protocol (HTTP) requests on the resource specified in the URL. REST also allows HTTP headers to be used to provide a request context around operations, so that the request can result in a desired media type specified in the header [16].

In the development of Onnikka, the thin-client MVC approach was implemented to separate the user interface and the business logic, and the RESTful design approach was followed to provide an easily accessible interface for performing operations on the application's resources. In Onnikka, each user-facing component is modeled as a resource, and each resource is identified by and resolved through a URL. These RESTful components may then internally implement the MVC pattern.

4.2 Implementation Considerations

The distribution of Onnikka's RESTful resources and their relationships (see Fig. 3) follows our suggested architecture design for persuasion context. As seen in Table 1, the *user context* component is implemented as the User resource. The *use context* component was implemented by an external content management system software component called RefineryCMS, which provided the environment to present the health information to the users and an interface for nontechnical personnel to update the information contents. The information itself was stored in the Week resource, corresponding to the *system-mediated messages* component. The Comment resource implements the *social interaction* component, and the *user-system interaction* component is realized as the Exercise entry, Food entry, Memo entry, and Weight entry resources.

Table 1. Implemented system resources according to the architectural components

Architectural component	**System resource**	**Function in Onnikka**
User context	User	Contains the user information.
Use context	RefineryCMS	Provides the user and admin interfaces.
User-system interaction	Exercise entry, Food entry, Memo entry, Weight entry	Contains the self-monitoring data submitted by the users.
Social interaction	Comment	Contains comments given in weekly discussions.
System-mediated messages	Week	Contains weekly articles, exercises and tips presented to users.

Each resource was designed utilizing the appropriate design principles suggested in the PSD model [8]. Principles from the primary task support category were used in the user context, use context, system-mediated messages, and user-system interaction related resources. In the use context and user-system interaction components, principles from the dialogue support category were also used, whereas the system-mediated messages component utilized system credibility support. The social interaction component utilized the social support principles. In more depth:

1. The User resource in the system represents the *user context*. It contains the information about the users required to provide them the appropriate content and persuasive features of the system. The basic information to authenticate the user is included, as well as the email address for sending persuasive messages via email reminders. The starting day of the intervention is stored to be able to present the user the appropriate use context according to their phase of intervention. The information about the tailoring preference for additional eating behavior content for the user is included. The target weight submitted by the user is stored to allow goal setting and personalized feedback on the weight loss progress. The user-submitted height is used to pass on to calculate users' body mass index (BMI) in their weight entries.
2. The external RefineryCMS content management system module provides the *use context*. It provides the interface for the administrators to submit the information contents into the system and generates the layouts to present the information to the users.
3. The *user-system interaction* is realized in the Weight entry, Memo entry, Exercise entry, and Food entry resources. These are used as tools for users to monitor their progress during the intervention, and receive feedback from the system. The persuasive guidelines utilized in the design of the components were mostly from the primary task support and dialogue support categories. The main persuasive feature here is self-monitoring. Personalized feedback based on the entries and the user information is given and additional eating behavior features are included for the tailoring-enabled users.
4. The *social interaction* component was actualized in the Comment resource. The users can leave comments on the week pages, where discussion related to the weekly subject is encouraged. The comment resource was designed primarily to support the persuasive guidelines of the social support category. Discussion among peers allows social learning, normative influence, and social facilitation within the system.
5. The *system-mediated messages* component was implemented as the Week resource. It includes the weekly introduction, main article, exercise, and tip provided by the health professionals. Each week may also contain an additional tailored article, exercise, and tip for the users with tailoring enabled. Each week contains the running number so that it is shown during the proper phase of intervention, and has a designated theme based on which it can be categorized in the indexing page of the weeks. The distribution of the information content into separate weekly portions implements the primary task support principles of reduction

and tunneling. The principles of system credibility support were also widely utilized—the informational content provided references different health authorities to improve the perceived trustworthiness, expertise, authority, and verifiability.

The software components presented above actualize the use, user, and technology contexts described earlier. This suggests that the proposed architecture can be used to implement a full-fledged BCSS.

5 Conclusion and Discussion

In this paper, we proposed a software architecture to be used in the development of behavior change support systems based on the analysis of the persuasion context presented in the PSD model. The proposed architecture was showcased through a weight loss and maintenance intervention system containing several persuasive features. The developed system seems to support the claim that the different aspects of the persuasion context can be successfully implemented in our architectural design.

The major implication of the research presented in this paper is that it provides a deeper view of the technological aspects involved in BCSS development. This will help in getting out of the black-box thinking approach to the research and development of persuasive systems, enabling designers to recognize characteristics of software components that may lead to the success or failure of applications. For practitioners, utilizing the architecture will help to create conventions that will ease and accelerate the development of new BCSSs.

One limitation of this study is that, since the developed system was in its early development phase, no perceived or actual persuasiveness has yet been measured and no expert evaluation on the implemented persuasive features has yet been conducted. Thus, the resulting system should be more carefully studied to indicate its impact. To further demonstrate the generalizability of the proposed architecture, multiple systems should be developed utilizing it. These should also include systems outside the currently used application domain area. For example, how the architecture would adapt to native mobile applications or web applications for sustainable energy behaviors could be studied.

Future work will include further definition and verification of our suggested architectural model. The developed system will be carefully evaluated with regards to various facets of persuasiveness, and its software functionalities will be studied with real users. In the future, we hope to be able to apply the suggested software architecture to multiple different problem domains and settings.

Acknowledgements. We wish to thank the contribution of the PrevMetSyn research consortium, especially Kreetta Askola, Heidi Enwald, Pasi Karppinen, Anna-Maria Keränen and Tuire Salonurmi.

References

1. 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)
2. Oinas-Kukkonen, H.: A foundation for the study of behavior change support systems. In: Personal and Ubiquitous Computing (2012) (Online First)
3. Lehto, T., Oinas-Kukkonen, H.: Persuasive Features in Six Weight Loss Websites: A Qualitative Evaluation. In: Ploug, T., Hasle, P., Oinas-Kukkonen, H. (eds.) PERSUASIVE 2010. LNCS, vol. 6137, pp. 162–173. Springer, Heidelberg (2010)
4. 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)
5. Lehto, T., Oinas-Kukkonen, H.: Persuasive Features in Web-Based Alcohol and Smoking Interventions: A Systematic Review of the Literature. Journal of Medical Internet Research 13(3), e46 (2011)
6. Chatterjee, S., Byun, J., Pottathil, A., Moore, M.N., Dutta, K., Xie, H(Q.): Persuasive Sensing: A Novel In-Home Monitoring Technology to Assist Elderly Adult Diabetic Patients. In: Bang, M., Ragnemalm, E.L. (eds.) PERSUASIVE 2012. LNCS, vol. 7284, pp. 31–42. Springer, Heidelberg (2012)
7. Mukhtar, H., Ali, A., Belaid, D., Sungyoung, L.: Persuasive Healthcare Self-Management in Intelligent Environments. In: 8th International Conference on Intelligent Environments, IE, pp. 190–197 (2012)
8. Oinas-Kukkonen, H., Harjumaa, M.: Persuasive systems design: Key issues, process model, and system features. Communications of the Association for Information Systems 24(1), 485–500 (2009)
9. Fogg, B.J.: Persuasive computers: Perspectives and research directions. In: Proc. of the SIGCHI Conference on Human Factors in Computing Systems, CHI, pp. 225–232 (1998)
10. Fielding, R.T., Taylor, R.N.: Principled design of the modern Web architecture. ACM Transactions on Internet Technology 2(2), 115–150 (2002)
11. Perry, D.E., Wolf, A.L.: Foundations for the study of software architecture. ACM SIGSOFT Software Engineering Notes 17(4), 40–52 (1992)
12. Enwald, H., Huotari, M.-L.: Preventing the Obesity Epidemic by Second Generation Tailored Health Communication: an Interdisciplinary Review. Journal of Medical Internet Research 12(2), e24 (2010)
13. Keränen, A.-M., Savolainen, M.J., Reponen, A.H., Kujari, M.L., Lindeman, S.M., Bloigu, R.S., Laitinen, J.H.: The effect of eating behavior on weight loss and maintenance during a lifestyle intervention. Preventive Medicine 49(1), 32–38 (2009)
14. Krasner, G., Pope, S.: A Description of the {Model-View-Controller} User Interface Paradigm in the Smalltalk-80 System. Journal of Object Oriented Programming 1(3), 26–49 (1988)
15. Leff, A., Rayfield, J.T.: Web-application development using the Model/View/Controller design pattern. In: Proceedings of the Fifth IEEE International Enterprise Distributed Object Computing Conference, EDOC 2001, pp. 118–127 (2001)
16. Battle, R., Benson, E.: Bridging the Semantic Web and Web 2.0 with Representational State Transfer (REST). Web Semantics: Science, Services and Agents on the World Wide Web 6(1) (2008)

Viewing and Controlling Personal Sensor Data: What Do Users Want?

Debjanee Barua[1], Judy Kay[1], and Cécile Paris[2]

[1] School of Information Technologies,
University of Sydney, Australia, 2006
dbarua@it.usyd.edu.au, judy.kay@sydney.edu.au
[2] CSIRO ICT Centre,
Cnr Vimiera and Pembroke Roads, Marsfield NSW 2122
cecile.paris@csiro.au

Abstract. Personal data from diverse sensors plays a key role in persuasive systems, especially those aiming to help people achieve long term goals. We need to gain an understanding of the ways people would like to capture and manage such data. We report the design and outcomes of a study exploring how people want to keep and control sensor data for long term health goals. We asked about three sensors, for weight, activity and sitting. We chose these for their diversity in terms of tracking progress on *means* and *end* goals, short and long term goals and differing sensitivity of the data. Our results show that people want to *use* and *control* a personal copy of such data and their preferences vary across different sensors. This points to the need for future persuasive systems to support these forms of user control over their sensor data.

1 Introduction

Persuasive computing has demonstrated the benefits to be achieved by computing systems that draw on psychological theories of behaviour change [6] to help people achieve long term goals, such as developing a healthy lifestyle or sustainable behaviour. Emerging sensors and mobile applications are key components of persuasive systems, because they enable people to easily capture fine grained personal data related to such goals. This has the promise of helping people achieve *long term* well-being goals, such as a healthy weight, high enough levels of activity and avoiding long periods of inactivity [8].

The long term nature of many important health goals poses key challenges because people need to achieve behaviour change and then maintain the new behaviour over the long term [13]. Important work has investigated aspects of the infrastructure for such technology for increasing activity [3,14,18], developing healthy eating [4,17] and sleep habits [11] and managing chronic diseases (e.g., diabetes [2]). To date, such work has focused on creating effective particular sensors and applications in persuading people. There has been little work on the issues of capture and management of the data associated with these, for example on data ownership [9]. Yet we urgently need to begin building persuasive systems in which personal sensing data is managed in line with people's wishes; for this

S. Berkovsky and J. Freyne (Eds.): PERSUASIVE 2013, LNCS 7822, pp. 15–26, 2013.

we first need to gain an understanding of whether people want to use such data in the long term and how they want to control it.

This paper reports the results of a questionnaire designed to gain insights into these issues in order to inform the design of systems that capture and manage long term personal sensing data. We explored whether people see value in having sensors automatically capture data about their weight, physical activity and inactivity. As such data is important for supporting long term persuasive systems, we investigated whether people see value in reflecting on such data over the short term as well as the *long term*, over months and years, and we analysed how this is affected by the nature of the sensor data. We then explored aspects of the *control* people want in the storage of such data.

2 Related Work

This section presents some key persuasive systems that made use of sensors and mobile phone applications to support people in achieving a healthy lifestyle. We focus on those systems where the evaluation involved long term studies and on the ways that this data was used and managed.

Lin and colleagues conducted a 14-week study for "Fish'n'Steps" [14], a game based intervention that gave both positive and negative feedback to encourage people to increase their daily physical activity. "WellnessDiary" [15,1] was a mobile application supporting behaviour change associated with managing well-being and chronic diseases. It was evaluated in two 12-week studies [16,15]. Subsequently, the researchers conducted the year-long trial "Nuadu trial", providing a platform that integrated several mobile applications and sensors. Similarly, the seminal work on "Ubifit" involved a three month study [3], assessing the *long term effectiveness* of their mobile sensing platform with its calm mobile display to encourage people become physically more active. These studies all share a focus in supporting people in health goals, and all involved long term user studies. However, none discussed the issues for long term storage and management of the personal sensor data.

There has been considerable work exploring people's privacy concerns. For example, a long term study with an exit interview provided insights into people's concerns about continuous recording and storing of their physical activity data [12], which included location, and particularly for some people, the capture of conversation. Some participants wanted their data stored in their own mobile devices, rather than a website. In our work, we are concerned with broader issues than privacy, as we want to reform our understanding into broad preferences for storing personal sensor information.

Other relevant persuasive research involves the models driving personalisation, based on analysing the personality of the users [7] and investigating their persuadability [10]. This relies on personal information about the user for the persuasion mechanisms. There has been acknowledgement of data ownership issues, and benefits of reuse of the data by multiple applications [9]. These pointed to the need for a greater understanding of how people want to be able to control this information, the goal of this work.

3 Study Design

We designed a survey to gain insights into the ways people want to control the personal sensor data at the core of many persuasive systems. We focused on the long term management of such personal data. We chose the context of health and wellness as it is important, understandable and valued by most people. We selected three forms of sensor data: weight, steps walked and time sitting (inactivity), because they are important [8,5] and because they represent different classes of personal health data.

Table 1. Hypotheses about use and control over long-term sensor data

H1.	People would like to have sensors automatically capture data about their weight, physical activity and inactivity.
H2.	People particularly value such sensor data for the *long term*.
H3.	The nature of the data affects H2.
H4.	People want their personal sensor data in storage they *control*.
H5.	The nature of the data affects H4.
H6.	Whether people have recorded such long term personal data affects H1-5.

Table 1 lists the hypotheses that drove the design of the questionnaire. H1 postulates that people would like to use personal data sensors. If people do see this as valuable, we would expect them to be willing to consider the long term management of the data collected. So the later hypotheses rely on this one. H2 deals with people's attitudes to *long term* collection of sensor data. This has not been previously explored in studies, but is important to know; if people do value long term data stores, it becomes urgent for mechanisms to ensure that such data is managed in a way that both is practical and meets their needs and preferences. H3 concerns the nature of different classes of sensor data; if this hypothesis is supported, that may mean that different storage and management mechanisms may be needed. H4 asserts that people want to control their sensor data, again an issue that has not been studied for persuasive systems but which has important implications for creating systems and their interfaces. H5, like H3, concerns the effect of different forms of data; if people have different preferences, these may call for different forms of support for managing and controlling the data. Finally, H6 asserts that people's previous practice will affect the earlier hypotheses. So, for example, if a person has logged their own personal data with existing tools over the long term, this may mean that their preference for managing such data with emerging technologies will be different from other people who have not logged such data.

The study was based on a questionnaire which introduced a series of sensing technologies, showing the ways that each could be used to capture data and presented it in a graph. It began by asking people if they could understand the graphs. Then it asked people about the usefulness of such information, followed by the management of the information. Finally there were questions about the management of the information.

We chose sensors that represent different classes of sensor data on several dimensions. Table 2 summarises these. The first form of data is weight, and the study illustrated this as captured with Withings scales[1]. We chose weight because many people consider it important, and all understand what it means and how bathroom scales measure it. The only new aspect for many people would be that the data is captured long term. This is an example of data that serves what has been described as *ends* goals [9], as distinct from *means* goals which are tackled to achieve an end; for example a person may have means goals associated with activity and food intake for the end goal of maintaining a healthy weight and, in turn, for the more general end goals of good health, feeling well and looking attractive. In addition, weight is an example of an attribute that changes slowly, over weeks or more. It is also an attribute about which we expected many people to be sensitive.

Table 2. Characteristics of sensors selected for the study

Attribute	**Device**	**Goal type**	**Reflection time scale**	**Sensitivity**
Weight	Withings	Ends	Weeks/months	High
Activity	Fitbit	Means	Day	Low
Sitting	Cushion	Means	Day	Medium

Our second sensor measures activity, and has quite different characteristics as indicated in the table. We presented it to participants as measured by a FitBit sensor[2]. As many people know about pedometers, we expected people to be able to readily appreciate what Fitbit does. We chose one more sensor, a hypothetical cushion that captures data about how long the person has been seated on it. This was chosen to give people a sense of the breadth of emerging sensor technology. The questionnaire also asked about gender, age, education and technical background, so that we could describe the participant population.

Table 3 shows the questions asked and how they were designed to test the hypotheses above and to assess a base hypothesis (H0) that participants understand graphs showing long term sensor data values. All asked the user to respond on a 5-point Likert scale, where 5 indicates strong agreement (SA) down to 1 for strong disagreement (SD). The code for each question in the first column will be used in the remainder of this paper. Q1 directly asks about the *usefulness* of the data. R1-3 explore preferences for use of the data over the *long term*. These inform H2, the comparison of responses across sensors informs H3, and the analysis in terms of previous recording of personal data informs H6. C1-6 ask about *control* based on having a copy of the data. All of these directly inform H4, and similarly to R1-3, the analysis across devices informs H5 and, in terms of self-logging, H6.

[1] `http://www.withings.com/en/bodyscale`
[2] `http://www.fitbit.com/`

Table 3. Summary of questions, mapped to hypotheses (H) from Table 1 The text shows questions for weight – other devices had similar questions

ID	Text of questions asked	H
Q0	I find this graph easy to understand.	H0
Q1	I would find it useful to have an easy way (like Withings) to record this sort of information about my weight.	H1
R1	I would usually review just the most recent value.	H2,3,6
R2	I would review trends over months.	H2,3,6
R3	I would review trends over years.	H2,3,6
C1	I would NOT want a copy as I do not see any need for it.	H4,5,6
C2	I would NOT want a copy as I feel that it is safer to have it kept by Withings/Fitbit/Cushion.	H4,5,6
C3	I would like a copy in case I find a use for it in the future.	H4,5,6
C4	I would like a copy so I can manipulate the data (e.g. in Excel).	H4,5,6
C5	I would want a copy as I am concerned about losing my data if I buy a new scale and delete my account.	H4,5,6
C6	I would want a copy as I am concerned about losing my data if the company goes out of business.	H4,5,6

The survey was conducted online with SurveyMonkey[3]. We sought participants in our university and others interested in personal informatics.

3.1 Participants

Fifty-four people completed the survey. Table 4 shows that almost 40% were women, about 40% aged 18-30, a similar proportion 31-45 and 17% were older. We have a bias towards highly educated people, with 78% having higher degrees, and the population is dominated by people with high technical skills. In summary, our study reflects the views of a highly educated and technically adept group of people. Table 5 shows the previous experience in recording the three forms of personal data. Notably, most people had never recorded this information, and 37.5% had never recorded even weight.

Table 4. Overview of participant demographics

Demographic	Number of participants
Gender	Male (33), Female (21)
Age	18-30 (23), 31-45 (22), >46 (9)
Education	High school(2), Undergraduate (10), Graduate (42)
Technical skill	Limited(2), Competent (7), Advanced (19), Expert (26)

4 Results

We first consider results in Figures 1 and 2 for H1-3, on perceptions of the value placed in capturing sensor data and on viewing it for very short term, months

[3] `http://www.surveymonkey.com/`

Table 5. Duration of recording personal information. Bold indicates the most common response for that data type.

Answer options	Weight	Steps	Inactivity
Never	**37.5%**	**66.7%**	**89.3%**
Less than 1 year	10.9%	19.3%	1.8%
1 to 5 years	23.4%	14.0%	5.4%
More than 5 years	28.1%	0.0%	3.6%

or years. Overall, 88% of participants found the graphs easy to understand (Q0) (mean score 4.18 overall, with device averages 4.15, 4.44 and 3.96 for Withings, FitBit and Cushion, respectively). This makes the rest of the responses meaningful. There was 76% agreement that sensor data is useful (mean scores, 3.92 overall, and 4.07, 3.89 and 3.78 per device). This supports H1.

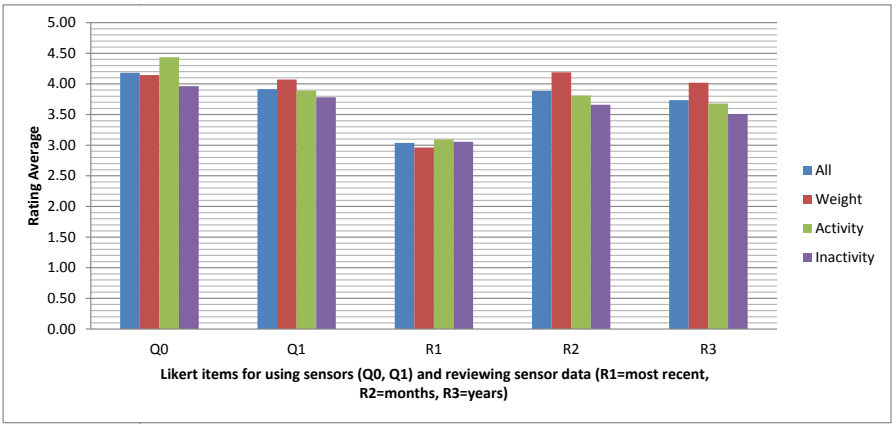


Fig. 1. H1-3. Average 5 point Likert scale responses. Q0 – graph understandable. Q1 – data useful. Would review data R1 (most recent), R2 (months) R3 (years). (N = 54).

Participants saw less value in short term data (overall mean scores: R1=3.03) than long term, months (overall mean scores: R2=3.9) and years (overall mean scores: R3=3.76). Kruskal-Wallis tests indicated significant differences in the response distributions across R1-3 for all three froms of data (χ^2=32.94, df=2, p=7.029e-08 for *weight*; χ^2=16.76, df=2, p=0.00023 for *activity* and χ^2=9.41, df=2, p=0.009 for *inactivity*).

This justified pairwise comparison which confirmed that participants valued *long term* use of the sensor data over *short term*. Mann-Whitney U tests showed strongly statistically significant differences between *R1 and R2* (Z=5.07, p=1.6e-07 for *weight*, Z=3.83, p=9.8e-05 for *activity* and Z=2.9, p=0.003 for *inactivity*) and *R1 and R3* (Z=-4.5, p=5.2e-06 for *weight*, Z=-3.1, p=0.002 for *activity* and Z=-2.23, p=0.02501 for *inactivity*). This supports H2.

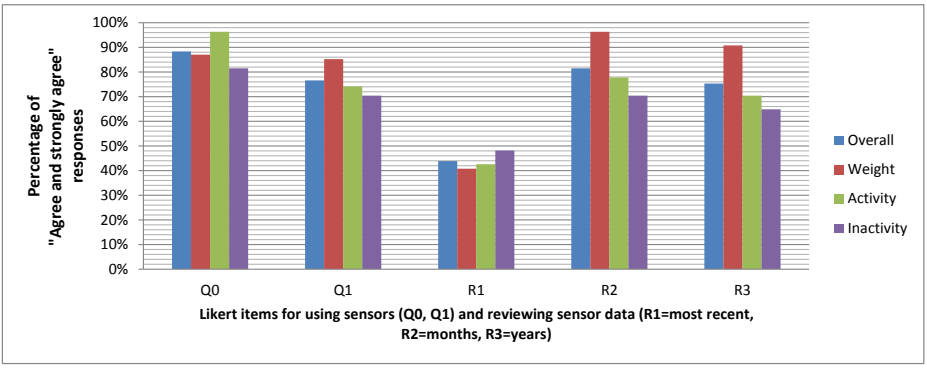


Fig. 2. Percentage agreement (Likert Scores 4 and 5) for questions in Table 1 (N = 54)

We repeated the analysis, excluding participants who indicated they did not want to use one or more of the sensor types. This left 42 participants. This yielded no significant differences for R1 across all forms of sensor data. However, Kruskal-Wallis analysis showed the means for R2 (χ^2=9.04, p<0.011) and R3 (χ^2=6.85, p<0.033) were significantly higher for *weight* than the others. Post-hoc analy-

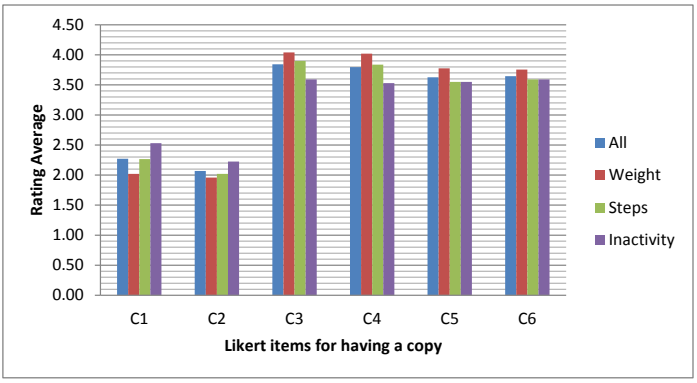


Fig. 3. H4,5. Average 5-point Likert scale responses. C1 – copy not needed. C2 – data safer with supplier. Reasons to want a copy: C3 – in case I find a use; C4 – to manipulate and analyse; C5 – in case lost by supplier; C6 – in case supplier fails. (N = 54).

sis using Mann-Whitney U tests showed significant differences between *weight* and *activity* (Z=2.6052, p<0.03) and *weight* and *inactivity* (Z=2.5886, p<0.02) for reviewing data over *months*. Similarly, for review over *years*, there were significant differences *weight* and *inactivity* (Z=-2.394, p<0.016) and marginally significant differences between *weight* and *activity* (Z=-1.7756, p<0.075). In both cases, there were no significant differences between activity and inactivity.

As the nature of weight data is different from the other two forms of data, the findings support H3.

Figures 3 and 4 show results about wanting a personal copy of data as a means of control (C1-C6 in Table 3). Just 16% of the 54 participants saw no need for a copy – C1 (mean=2.24) and 6% agreed the data is safer with the manufacturer – C2 (6%, mean=2.0).

Over 70% agreed with C3 (mean=3.82), wanting a copy for the future, and 65% C4 (mean=3.9) that this was to manipulate the data. Nearly 60% of participants were concerned about losing their data (C5 mean=3.58, C6 mean=3.54). This supports H4, on wanting a copy.

For H5, we compared the value placed on having a copy, across the devices (N=42 as in H3) with marginally significant differences for C1 (χ^2=4.99, df=2, p=0.082), C3 (χ^2=5.43, df=2, p=0.066) and C4 (χ^2=6.22, df=2, p=0.045). A pairwise comparison showed statistically significant differences between *weight* and *inactivity* (Z=2.6052, p=0.009183 for C1, Z=2.3259 p=0.02002 for C3 and Z=-2.0718, p=0.03829 for C4) and *weight* and *activity* (Z=1.913, p=0.05574 for C1 only). This supports hypothesis H5, with weight again being different from the others.

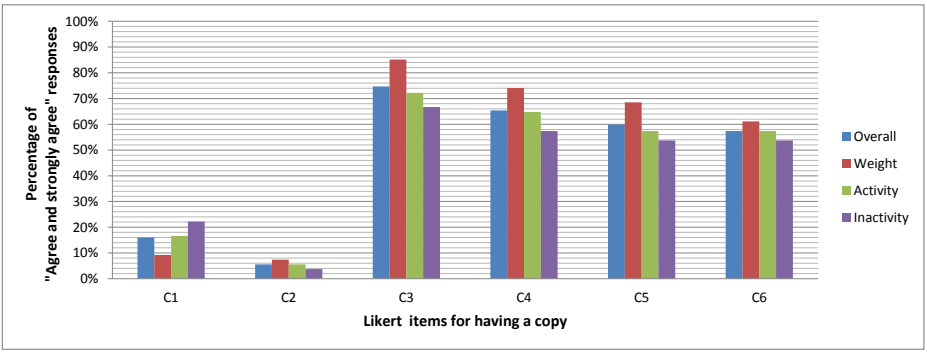


Fig. 4. Percentage agreement (Likert Scores 4 and 5) for questions in Table 3 (N = 54)

For Hypothesis H6, we compared answers for participants based on their previous recording of weight or activity. (We ignored inactivity as so few had recorded it – <10%.) For weight, the groups were: "Group 1" (N=18; no recording), "Group 2" (N=19; recorded, but for <= 5 years) and "Group 3" (N=18; > 5 years). A Friedman test for R1-R3 and C1-C6 showed no statistically significant differences. For activity, we compared two groups: "Group A" (N=27; no recording) and Group B (N=15; < 0 <= 5 years). None had recorded steps for > 5 years. On a Wilcox test, the differences were statistically significant for R1 (Z=-2.5057, p-value=0.01208), R2 (Z = 0.8425, p-value = 0.04021) and, R3 (Z=1.5839, p-value=0.0982) but not for C1-C6. This refutes hypothesis H6.

5 Discussion

Table 6 summarises our results and implications for designing persuasive systems that capture sensor data to make it available to users. H1 (Row 1) confirms that people see value in capturing sensor data and to see it in graphs like those we presented. This is the first study we know of to confirm the rarely stated assumptions of many ubicomp persuasive projects and products similar to those presented to our participants. This question was a preliminary to the later ones – perceptions about data control are meaningful only if people consider that data useful.

Table 6. Design implications from findings on hypotheses in Table 1

H	Support?	Design Implications
1	Yes	People perceive value in health data from sensors – confirms assumptions of many research projects and commercial products
2	Yes	People value access to their sensor data over the long term, *months* and *years* – systems capturing such data need to hold it, and make it available for inspection by the user, over these time periods.
3	Yes	People particularly value long term access to data about their weight, the slow-to-change *ends*-goal, compared with the daily-changing *means* goals of activity and inactivity – point to the potential need for different duration for the storage of different data forms.
4	Yes	People want a copy of their own data – this relates to issues of data ownership – we need mechanisms enabling people to *easily* get such a copy, more than just an API.
5	Yes	A copy of weight was significantly more important. Overall, weakly significant differences across sensors on seeing a need for the data, 'in case I find a use', with significance for being able to manipulate it – people particularly want a copy for some sensor data.
6	No	H4 holds whether people previously tracked themselves, or not.

H2 explored the value placed on short versus long term data. Participants clearly saw more value over the long term. This has important implications for the design of persuasive systems that capture such data. Research projects, with their limited timespans, tend to ignore these issues. Practical long term deployed systems will need to deal with it.

H3 explored differences in H2 across the three sensors. Participants clearly saw more importance in weight than the others, with more than >90% keen to review long-term weight, >70% activity and >60% sitting. It may be that weight, as an *ends*-goal, is inherently more interesting in the long-term. The displays we showed participants favour thinking about long-term trends. They do not even hint at options like ambient or glanceable displays. Equally, it may be that this is due to people's familiarity with weight, and tracking it and, some made comments indicating they cared more about it. With our focus on data ownership preferences and implications for building persuasive systems, our key conclusion is the different value people placed in weight points to the need for

different storage duration according to the sensor type. Put differently, people may want some of their sensor data kept over the very long term, and other data could be forgotten by persuasive systems.

We now consider the issues of having a personal copy of one's sensor data, (Rows 4 and 5 of Table 6). Our participants clearly wanted a personal copy of their sensor data. This is currently not well supported in persuasive systems. Systems must provide highly usable mechanisms for copying. An API is a good start, but it is not sufficient. The strength of participant views is eloquently expressed in their free comments. For example, one commented: *I agree with all of the use cases for why someone would want their data. I think it's just nice to have ownership over your data since you are volunteering it to them. It is personal information, after all.*

Another participant explained how they wanted to manipulate their own copy of weight data: "*The most important factor here for me would be wanting to manipulate the data in some way in, for example, Excel. Weight data ... is not actually a very relevant measure by itself, and I might want to combine it with a waist or fat caliper measurement which would allow me to estimate body fat % (a more useful measure). I am also interested in power-lifting so I may want to combine it with values for my 1 rep max on certain lifts to track my strength-weight ratio.*"

H5 was supported, with participants particularly wanting a copy of weight data. While participants wanted a copy of all sensor data, this was most important for weight, our slow-changing *ends* goal. This makes it important that persuasive systems make such data easy to copy.

H6 explored if previous experience in recording data affected perceptions of all aspects above. We wondered if people who had taken the trouble to capture data, especially over the long term, had different views. This was not supported. This suggests that our results hold for a broad population, including those who have tracked themselves, and those who have not.

6 Conclusions

Personal sensor data is critical for important classes of persuasive systems. This makes it important to ensure that it will be managed as people wish. No previous study explicitly explored if people want to keep long term personal sensor data and control it, based on having a copy in storage that they control. Our study addresses these issues for a set of three, quite different sensors. They represent classes of *means* and *ends*-goal data capture, corresponding to fast-changing aspects of activity and inactivity that people can decide to alter in the short-term and slow-changing weight which can be tracked.

Our analyses of responses from 54 participants support our hypotheses that people value such data, though this does not apply uniformly across sensors. People want a copy of their sensor data, even if they do not yet want to manipulate it; but in case they find a use for it. With the state of present sensors, this is a problem. Typically, each sensor, and its associated data, is under the

control of its manufacturer. While they may provide an API, this does not make it feasible for most people to get a copy of their own data. This points to the need to design persuasive systems that enable people ensure that their sensor data can be captured over the long term, that some data can be forgotten by systems but other data should be remembered, and there should be mechanisms for people to easily get a personal copy. Many important issues remain to be explored, notably how to ensure people can make effective use of their personal copy of data.

This is a first exploration of user preferences for managing data from personal sensors for persuasive systems. It has limitations. It considered just three sensors; this avoided making the survey too long and tedious. To mitigate this, we chose the sensors carefully, to represent diverse forms of data, over a range of important aspects of health. Our participant population is distinctive; they are highly educated and technically competent. We approached such people as the survey demanded thinking about complex and unfamiliar ideas and technical concerns. We have no reason to suppose that this population would have cause to put different value on data about their health than a broader population. Our goal was to inform the design of persuasive systems that involve long term personal data, particularly about health. Overall, our study points to the need for infrastructures and mechanisms that make it easy for people to automatically capture, control and review their data over the long term.

Acknowledgement. This work was supported by funds from CSIRO and Australian Research Council Grant DP0877665.

References

1. Ahtinen, A., Mattila, E., Vaatanen, A., Hynninen, L., Salminen, J., Koskinen, E., Laine, K.: User experiences of mobile wellness applications in health promotion: User study of wellness diary, mobile coach and selfrelax. In: Pervasive Health, pp. 1–8 (2009)
2. Chatterjee, S., Byun, J., Pottathil, A., Moore, M.N., Dutta, K., Xie, H(Q.): Persuasive Sensing: A Novel In-Home Monitoring Technology to Assist Elderly Adult Diabetic Patients. In: Bang, M., Ragnemalm, E.L. (eds.) PERSUASIVE 2012. LNCS, vol. 7284, pp. 31–42. Springer, Heidelberg (2012)
3. Consolvo, S., Klasnja, P., McDonald, D.W., Avrahami, D., Froehlich, J., LeGrand, L., Libby, R., Mosher, K., Landay, J.A.: Flowers or a robot army?: encouraging awareness & activity with personal, mobile displays. In: Ubicomp, pp. 54–63 (2008)
4. Denning, T., Andrew, A., Chaudhri, R., Hartung, C., Lester, J., Borriello, G., Duncan, G.: Balance: towards a usable pervasive wellness application with accurate activity inference. In: HotMobile 2009, pp. 5:1–5:6 (2009)
5. Dunstan, D., Kingwell, B., Larsen, R., Healy, G., Cerin, E., Hamilton, M., Shaw, J., Bertovic, D., Zimmet, P., Salmon, J., et al.: Breaking up prolonged sitting reduces postprandial glucose and insulin responses. Diabetes Care 35(5), 976–983 (2012)
6. Fogg, B.: A behavior model for persuasive design. In: Persuasive 2009, pp. 40:1–40:7. ACM (2009)

7. Halko, S., Kientz, J.A.: Personality and Persuasive Technology: An Exploratory Study on Health-Promoting Mobile Applications. In: Ploug, T., Hasle, P., Oinas-Kukkonen, H. (eds.) PERSUASIVE 2010. LNCS, vol. 6137, pp. 150–161. Springer, Heidelberg (2010)
8. Haskell, W., Lee, I., Pate, R., Powell, K., Blair, S., Franklin, B., Macera, C., Heath, G., Thompson, P., Bauman, A.: Physical activity and public health: updated recommendation for adults from the american college of sports medicine and the american heart association. Medicine and Science in Sports and Exercise 39(8), 1423 (2007)
9. 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)
10. 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)
11. Kim, S., Kientz, J.A., Patel, S.N., Abowd, G.D.: Are you sleeping?: sharing portrayed sleeping status within a social network. In: CSCW, pp. 619–628 (2008)
12. Klasnja, P., Consolvo, S., Choudhury, T., Beckwith, R., Hightower, J.: Exploring Privacy Concerns about Personal Sensing. In: Tokuda, H., Beigl, M., Friday, A., Brush, A.J.B., Tobe, Y. (eds.) Pervasive 2009. LNCS, vol. 5538, pp. 176–183. Springer, Heidelberg (2009)
13. Klasnja, P., Consolvo, S., Pratt, W.: How to evaluate technologies for health behavior change in HCI research. In: CHI 2011, pp. 3063–3072. ACM (2011)
14. 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)
15. Mattila, E., Korhonen, I., Salminen, J., Ahtinen, A., Koskinen, E., Sarela, A., Parkka, J., Lappalainen, R.: Empowering citizens for well-being and chronic disease management with wellness diary. IEEE Trans. IT in Biomed. 14(2), 456–463 (2010)
16. Mattila, E., Parkka, J., Hermersdorf, M., Kaasinen, J., Vainio, J., Samposalo, K., Merilahti, J., Kolari, J., Kulju, M., Lappalainen, R., Korhonen, I.: Mobile diary for wellness management – results on usage and usability in two user studies. IEEE Trans. IT in Biomed. 12(4), 501–512 (2008)
17. Orji, R., Mandryk, R.L., Vassileva, J.: Towards a Data-Driven Approach to Intervention Design: A Predictive Path Model of Healthy Eating Determinants. In: Bang, M., Ragnemalm, E.L. (eds.) PERSUASIVE 2012. LNCS, vol. 7284, pp. 203–214. Springer, Heidelberg (2012)
18. Purpura, S., Schwanda, V., Williams, K., Stubler, W., Sengers, P.: Fit4life: The design of a persuasive technology promoting healthy behavior and ideal weight. In: CHI, pp. 423–432 (2011)

Colours That Move You: Persuasive Ambient Activity Displays

Patrick Burns[1], Christopher Lueg[1], and Shlomo Berkovsky[2,*]

[1] School of Computing and Information Systems, University of Tasmania, Hobart, Australia
{firstname.lastname}@utas.edu.au
[2] NICTA, Sydney, Australia
Shlomo.Berkovsky@nicta.com.au

Abstract. Regular physical activity is essential for maintaining good health. Unfortunately due to the increasingly sedentary nature of modern life many people are not active enough. Although most have the ability to be more active they lack sufficient motivation. Persuasive technology could help to address this problem. We discuss the use of persuasive ambient displays, specifically wearable ambient displays, to motivate users to be more active. We show that it is critical to carefully consider how best to visualise data with a display in order to realise maximum persuasive effect. We present as a case study our ongoing design and development of ActivMON, a wearable persuasive ambient display.

1 Introduction

Regular physical activity is essential to maintaining a healthy body and mind. Unfortunately due to increasingly sedentary lifestyles many people are not active enough. The subsequent rise in the rate of overweight and obesity has resulted in a concomitant rise in the prevalence of certain chronic diseases such as type 2 diabetes and cardiovascular diseases [1]. Technology could play a role in addressing this problem by persuading or motivating users to engage in and sustain regular physical activity.

Of the different approaches to using technology to encourage physical activity, each varies in terms of the level of persuasion employed. Technologies using an activity tracking or information-based approach, such as pedometers or smartphone activity trackers, persuade implicitly on the basis that self-monitoring can encourage behaviour change [2]. Technologies such as exergames persuade implicitly in that physical activity is linked to an enjoyable or interesting activity – gaming. Technologies using motivational messages or virtual coaching employ more direct persuasion, operating in the role of a social actor [3].

Fogg's Behavior Model postulates that motivation, ability and a trigger must converge at the same moment for behaviour to occur [4]. Healthy adults would place moderate to high on the scale of ability to do exercise. Although there are barriers such as lack of time, lack of money (to join a gym or take up a sport), or the weather, most people could find some time during their day to do some sort of physical activity

[*] The review of this paper was managed by Jill Freyne.

S. Berkovsky and J. Freyne (Eds.): PERSUASIVE 2013, LNCS 7822, pp. 27–32, 2013.

if they wanted to. Arguably the central reason people fail to become more active and to pass Fogg's "activation threshold" is a lack of motivation, compounded to an extent by a lack of appropriate triggers. Therefore it is sensible that persuasive exercise technologies attempt to motivate the user in some fashion and provide a behavioural trigger.

The same model could be applied to users' willingness to engage with activity motivating technologies. The complexity and usability of a technology affects users' ability to engage with it. The perceived usefulness of a technology affects users' motivation to engage with it. With reference to Fogg's parabolic activation threshold [4], users would need to be highly motivated to engage with a technology that was hard to use. A simpler technology would cater to less motivated users.

Ubiquitous computing technologies, and specifically ambient displays, are an example of a technology that is intuitive and easy to engage with [5]. There are a number of examples in the literature: Lin et al. [6] represented users' state of activity using virtual fish in a fish tank – users could see if they were active enough from the size of their fish. Consolvo et al. [7] used a garden metaphor where the appearance of flowers and butterflies represented activities performed. Rogers et al. [8] used abstract artwork in a building to represent the occupants' stair use behaviour. Lim et al. [9] developed a display for a user's shoes that lit when they walked.

Expanding on the above works, we developed ActivMON – a wrist-based ambient display that detects users' activity using a motion sensor and represents this activity with a coloured light (Fig. 1). The light is red at the start of each day (no activity performed) and changes on a continuous spectrum through yellow to green as the user performs activity toward a daily goal.

Fig. 1. ActivMON ambient display

Fig. 2 illustrates the ActivMON visualisation. At the end of each week the user's historical activity level *h* is calculated as an average of each day's activity over the past week. A new daily goal for the coming week is set at 10% above *h*, shown as goal line *g*. On the y axis *c* represents the user's current activity level. At the beginning of each day *c* would be reset to zero resulting in a red visualization. As the current activity level *c* increased toward the goal *g* the light colour would change from red through yellow to green. We considered having users set their own goals but felt that it was more interesting to explore the possibility of adaptive goal setting, particularly given the desire to maintain a simple interface and minimise user burden.

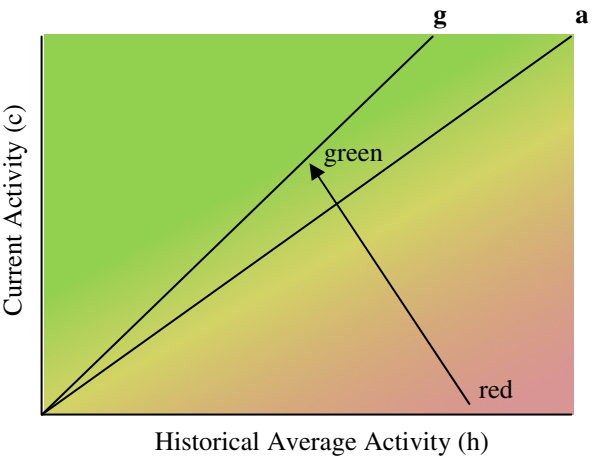


Fig. 2. ActivMON activity to colour visualisation

We evaluated ActivMON in a pilot study with five users over two weeks (results reported in [10]) and discovered that its persuasive or dissuasive effects were closely linked to the operation of the chosen visualisation. Users said the red-to-green visualisation was intuitive – that "red is bad and green is good" – but users' response to the visualisation varied. One user said it was like "climbing the hill each day". Another said that the natural 24-hour cycle of the device was not ideal, as some people do more or less activity each day over the course of a week. Users reported that on some days it was too easy to make the device turn green and that they were no longer motivated when they reached their daily goal.

These problems were due to our approach of treating each day's physical activity as a discrete unit. Goals were based on daily activity and only activity over a single day was shown on the display. We found that users' day to day activity levels are volatile and were too sensitive a measure to display directly. However, the current level of physical activity can also be represented as a continuum over a longer period of time with rises and falls in activity. A continuum visualisation would be more stable, allowing users to focus on time variant trends rather than short term activity fluctuations. We present this new visualisation and argue that it could be applied to represent any recurring activity or behaviour.

2 A Continuum Approach

Rather than visualising cumulative activity over the course of a day a continuum display would show average daily activity over a sliding window of some period of time, such as the past seven days. Goals would continue to be set at the end of each week, however rather than trying to reach a goal every day users would try to maintain an average activity level as close to the goal as possible. Users would need

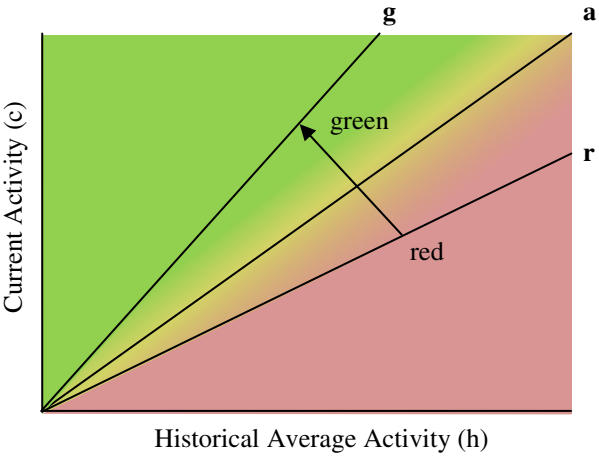


Fig. 3. Continuum colour to activity visualisation

to maintain a consistent level of activity to keep the light green. Inactivity would cause the light to slide backwards into yellow and red.

Fig. 3 illustrates the continuum visualisation. As before at the end of each week the user's historical activity level h is calculated as an average of daily activity over the past week. A new daily goal for the coming week g is set at 10% above h. However, current activity c is no longer reset to zero at the beginning of each day. Instead c is calculated as an average of current and past daily activity over a seven day sliding window. The region over which the light changes from red to green no longer starts at c=0 but instead is bounded at the green end by the goal line g and at the red end by a new "red line" r. As the user's current activity level c approaches g the light will turn green. If c drops to a and then to r the light will change through yellow to red.

As an example, imagine that we have already collected a week's worth of activity data from a user. With reference to the continuum shown in Fig. 4, let her average activity level over the past week be a_1. Her goal g_1 is to try over the coming week to sustain a new average activity level greater than the previous week. At the start of the week her current average activity level over a seven day sliding window will equal her previous week's historical average ($c = a_1$) and she will receive a yellow indication. Assume that she is very active on the first day of the week. Her current average activity level will rise toward g_1 and the ambient visualisation will change toward green. Assume that she then becomes busy with work the next day and is very inactive. Her current average activity level will drop back toward a_1 and she will see a yellow colour. If she is inactive the next day as well her activity level will drop further and as she approaches r_1 she will see a red colour. She may then compensate with more activity over the next few days and move her activity level back toward g_1 and therefore into the green zone.

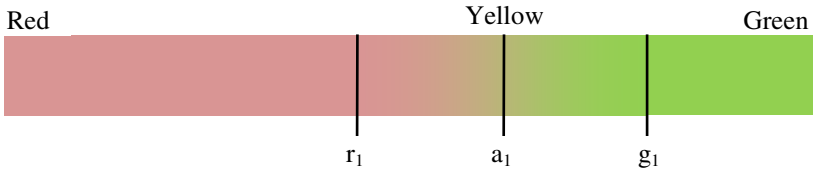


Fig. 4. Example red to green visualisation

By treating physical activity as a continuum rather than a series of discrete events we are able to create a visualisation that smooths over the natural variability of day-to-day activity. This smoothed visualisation should allow users to focus on trends and consistency rather than short-term changes.

The distance of the goal and red lines *g* and *r* from the average line *a* will affect the perceived difficulty of achieving the set goal as well as the leniency of the system to recidivism. The closer *g* is to *a* the easier the goal will be to achieve. The closer *r* is to *a* the less leniency there will be when the user has one or two inactive days in a week. These variables may affect the persuasiveness of the visualisation to the user. As was found in our preliminary evaluation, a goal (getting to the green zone) that is too easy to achieve provides little challenge, but a goal that is too difficult can be de-motivating (the light is always yellow or red). A system that is not lenient enough (where one day of sedentary behaviour produces a red indication) could be disheartening whereas a system that is too lenient might provide little incentive to maintain a consistent level of activity [11].

We intend to undertake a further user study to determine optimum values for each of the variables in the visualisation. That is, the way in which goals are set based on average activity (using a constant or proportional increase) and the placement of the "red line". We will create a number of variations and test each with a small group of users, assessing the persuasiveness of each variation through post-study questionnaires and semi-structured interviews.

3 Conclusions

We argue that Fogg's Behavioral Model provides an answer not only to the question of why people are not more physically active, but also to the question of which types of technology might be more effective in persuading people to be more active. We suggest a focus on technology that makes activity information simple to engage with, for example ambient displays.

We discussed the results of our initial evaluation of an ambient exercise display where we found that the ambient visualisation used was closely linked to the persuasive or dissuasive effect of the display on users. We presented a new "continuum" visualisation design for individual activity that we have developed in response to user feedback. This continuum approach could apply to visualisation of any recurring activity that can be represented on a linear scale, such as frequency of drinking, smoking or fast food consumption.

It is clear that careful attention to design is warranted to avoid creating a visualisation that could be ineffective, or worse dissuasive to users. We intend to conduct further user studies to ascertain the optimal settings for the continuum visualisation with a view to integrating it into a future wearable ambient display.

Acknowledgments. The support of the Tasmanian ICT Centre, through the CSIRO – University of Tasmania Research Scholarships, is acknowledged. The Tasmanian ICT Centre is jointly funded by the Australian Government through the Intelligent Island Program and the CSIRO. The Intelligent Island Program is administered by the Tasmanian Department of Economic Development.

References

1. Fletcher, G.F., Balady, G., Blair, S.N., Blumenthal, J., Caspersen, C., Chaitman, B., Pollock, M.L.: Statement on exercise: benefits and recommendations for physical activity programs for all Americans: a statement for health professionals by the Committee on Exercise and Cardiac Rehabilitation of the Council on Clinical Cardiology. Circulation 94, 857–862 (1996)
2. Kopp, J.: Self-monitoring: A literature review of research and practice. Social Work Research & Abstracts 24(4), 8–20 (1988)
3. Fogg, B.J.: Persuasive Technology: using computers to change what we think and do. Ubiquity (December 2002)
4. Fogg, B.J.: A behavior model for persuasive design. In: Persuasive 2009. ACM Press, New York (2009)
5. Weiser, M., Seely Brown, J.: Designing calm technology. PowerGrid Journal 1(1), 75–85 (1996)
6. 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)
7. Consolvo, S., Klasnja, P., McDonald, D.W., Avrahami, D., Froehlich, J., LeGrand, L., Libby, R., Mosher, K., Landay, J.A.: Flowers or a robot army? encouraging awareness & activity with personal, mobile displays. In: Proc. Ubicomp 2008, pp. 54–63. ACM Press (2008)
8. Rogers, Y., Hazelwood, W.R., Marshall, P., Dalton, N., Hertrich, S.: Ambient influence: can twinkly lights lure and abstract representations trigger behavioral change? In: Proc. UbiComp 2010, pp. 261–270. ACM Press (2010)
9. Lim, B.Y., Schick, A., Harrison, C., Hudson, S.E.: Pediluma: motivating physical activity through contextual information and social influence. In: Proc. TEI 2011, pp. 173–180. ACM Press (2011)
10. Burns, P., Lueg, C., Berkovsky, S.: ActivMON: Encouraging Physical Activity through Ambient Social Awareness. In: Proc. CHI EA 2012, pp. 2363–2368. ACM Press (2012)
11. Sweetser, P., Wyeth, P.: GameFlow: a Model for Evaluating Player Enjoyment in Games. Computers in Entertainment 3(3) (2005)

Personalized Persuasive Technology – Development and Validation of Scales for Measuring Persuadability

Marc Busch[1], Johann Schrammel[1], and Manfred Tscheligi[1,2]

[1] CURE – Center for Usability Research & Engineering, Vienna, Austria
(busch,schrammel,tscheligi)@cure.at
[2] University of Salzburg, ICT&S Center, Salzburg, Austria
manfred.tschligi@sbg.ac.at

Abstract. In this study we develop and validate an inventory for measuring persuadability to selected persuasive strategies. The development of the initial inventory was successful by means of internal consistency and item-scale correlation for the persuasive strategies *rewards, competition, social comparison, trustworthiness* and *social learning*. The inventory can be used to estimate susceptibility to persuasive strategies to personalize persuasive technology according to the users' personality based on self-reports. This can help system designers to make informed design decisions and to adapt persuasive technology.

1 Introduction

Persuasive Technology has been in the focus of HCI for over a decade – getting popular with Fogg's [1] *Persuasive Technology: Using Computers to Change What We Think and Do* in 2002. Different persuasive strategies such as e.g. *rewards* or *social comparison* have been suggested to be used in such technology. These persuasive strategies are supposed to change human attitudes and behavior. Innovative methods have been used to evaluate the effect of these persuasive technologies [2]. However, there are individual differences in the susceptibility to different persuasive strategies – called persuadability [3][4][5][6].

To create personalized persuasive technology (which is expected to have greater impact than not-personalized technology), it is necessary to be able to estimate the susceptibility of a person to different persuasive strategies (persuadability) - this is also referred to as "Persuasion Profiling" [4]. As a contribution to improve the design of persuasive technology, we develop and initially validate scales to measure *persuadability* for selected *persuasive strategies* for which specific psychometric inventories do not exist yet, namely: *Rewards, competition, social comparison, trustworthiness, simulation, reduction* and *social learning*, based on a collection by Torning and Oinas-Kukkonen [7]. The initial scale development was successful (by means of internal consistency and item-scale correlation) for all scales except for: *simulation* and *reduction.*

S. Berkovsky and J. Freyne (Eds.): PERSUASIVE 2013, LNCS 7822, pp. 33–38, 2013.

2 Related Work: Persuasive Technology, Persuadability and Scale Development

A broad range of persuasive strategies have been identified, the most prominent are from Fogg (7 strategies) [1], Cialdini (6 strategies) [8] and a collection by Torning and Oinas-Kukkonen (28 strategies) [7]. Persuasive Technology is technology which incorporates one or more persuasive strategies. People differ in their susceptibility to different persuasive strategies. This leads to the assumption that personalized persuasive technology is more successful than persuasive technology that is intended to "fit for all" [4].

Persuadability in an HCI-context was introduced by Kaptein et al. [3] in 2009 and examined in several studies (see for example [4][5][9]). It is defined as *the individuals' susceptibility to persuasive strategies and principles*. Earlier, in the 1970's, Bauer defined persuadability in a marketing context as *the disposition to shift one's attitude or judgment after being exposed to a persuasive communication* [10]. Focused on persuasive systems, Kaptein, Lacroix and Saini [9] defined in 2010 persuadability as a *tendency to comply to implementations of persuasive strategies*. Several quantitative measures for the susceptibility to *persuasive strategies* have been developed, either in the domains of communication, cognition and social psychology (here called *persuasibility* instead of *persuadability*) [11][12] or in the domain of *persuasion and HCI*. Especially the work of Kaptein and colleagues (e.g. [3][4][5][6][9]) is of relevance. They developed a 7-Item *persuadability* questionnaire measuring persuadability to the persuasive strategies *scarcity* and *authority* [3], and further a *persuadability* questionnaire with six sub-scales *reciprocity, scarcity, authority, commitment, consensus, liking,* available as 12 and 32-item version [4]. The questionnaire items are statements that contain indicators for the susceptibility to single persuasive strategies, e.g. "*When I am in a new situation, I look at others to see what I should do*" focusing on the persuasive strategy *consensus* [4].

3 Development and Validation of Persuadability Scales

The goal of this work is to develop a reliable self-report inventory for important persuasive strategies for which no such questionnaires exist yet. In detail, the strategies *rewards, competition, social comparison, trustworthiness, simulation, reduction* and *social learning* are considered (chosen by experts – see next section). For scale development, we follow selected steps from a process that is proposed for construct measurement and validation procedures in management information systems: Definition of the construct, the generation of items to assess the construct (in section: 3.1 First Activity: Creation of the Scales), an assessment of content validity of the items, the collection of data to conduct a pretest and the scale purification and refinement (in section: 3.2 Second Activity: Initial Validation of the Scales) [13].

3.1 First Activity: Creation of the Scales

This section covers the first part of the scale development process [13], the item generation, and an initial assessment of the content validity of the items. We follow the best practice example of the development of the *AttrakDiff*-questionnaire [14] and conduct an HCI-expert workshop for the creation and expert validation of initial scales with 6 experts and a moderator. First, participants were introduced into concepts and definitions of persuasion and persuasive strategies in ICT and also into the concept of persuadability as (more or less) stable traits of personality that can predict the individuals' susceptibility to persuasive strategies. Then, the experts were presented the definitions of 28 persuasive strategies [7] (out of Torning & Oinas-Kukkonen's collection) together with persuasive cues that can be implemented in interactive systems. As persuasive strategies are overall very abstract concepts and as some persuasive strategies are more abstract than others, participants had the task to choose a number of appropriate persuasive strategies as a basis for the creation of persuadability scales that fulfill the following criteria: Persuasive strategies can *reasonably* be translated into persuasive cues, these persuasive cues have nearly the same "estimated effect" on people and the susceptibility to these persuasive cues (persuadability) can be measured in a *meaningful* way by quantitative, verbal self-assessments. Per voting, experts chose 8 from 28 persuasive strategies that fulfilled the criteria best to be transferred into a questionnaire.

In a next step, the experts had to formulate self-assessment questionnaire items (in german) to assess the persuadability for the 8 chosen persuasive strategies. Experts were asked to develop precise items and to cover only one aspect per item. Overall, experts formulated 15 items for the scale *simulation*, 13 items for *trustworthiness*, 12 items for *rewards*, 12 items for *social comparison*, 11 items for *social learning*, 9 items for *competition*, 7 items for *reduction* and 3 items for *recognition*.

Then, an expert validation of the items was undertaken with the same experts that created the scales: Each expert had to rate each item of the 8 scales by means of fulfillment of the following criteria: If the item fits to the focal definition of the persuasive strategy dimension (content validity), if it covers only one specific aspect and if it can be assumed that the item represents an unidimensional scale together with all other items of that scale. Experts had to rate on a dichotomous scale: Fulfillment of the criteria (+1) or no fulfillment of the citeria (-1). In case of uncertainty, items did not have to be rated (0). For each scale, the 6 best rated items were chosen for a first version of the persuadability-inventory. As for the scale *recognition* only 3 items had been formulated, this scale was excluded from the inventory. The outcome of these steps is a first version of the persuadability-inventory with 7 scales with overall 42 items (see Table 1; all items).

3.2 Second Activity: Initial Validation of the Scales

This section describes the second part of the scale development process [13], a first item analysis (reliability by means of internal consistency and item-scale correlations)

and the refinement of the scales. The items were administered in an online-study. First, participants were shortly introduced into the topic of the study and then administered the first version of the persuadability-inventory (see Table 1; all items) with a randomized order (to avoid an item-order bias). Answer format was a nine-point rating scale, ranging from 9= *Fully agree* to 1= *Fully disagree*. Participants were asked for sex, age and education.

We analyzed data from n=167 participants (49,1% male; mean age: M= 37,6; SD= 14,4; education levels: 6,6% secondary school, 15,6% apprenticeship, 41,3% A-Levels, 8,4% bachelor's degree, 26,3% master's degree, 1,8% PhD level). For estimating reliability (internal consistency) we calculated *Cronbachs Alpha*. Also, we calculated corrected item-scale-correlations (ISC). All internal consistencies can be seen as sufficiently high to ensure reliability (over or near 0.7), except for *simulation* and *reduction*. Additionally, all ISCs of these two scales are under the commonly accepted critical value of 0.3. As a consequence, these two scales were skipped from the questionnaire. From the remaining scales, we eliminated all items with ISCs under 0.3. Although the *Cronbachs Alpha* of the *trustworthiness*-scale is below 0.7 (.472), its internal consistency is likely to increase after elimination of the items with ISCs under 0.3 (this new internal consistency has to be estimated with a different sample – it is not appropriate to estimate it again with the same sample). The outcome of these steps is an iterated second version of the persuadability-inventory with 5 scales and overall 25 items (see Table 1; removed items are greyed out, items indicated with (r) have to be reversed before calculating an overall score). Scores resulting from questionnaire can be interpreted like this: Participants having higher scores in one or more of the scales are expected to be more susceptible to these specific persuasive strategies.

4 Discussion and Future Work of the Scale Development

This paper has shown the initial development (creation, expert validation, item analyis by means of internal consistency in online study) of 5 scales for measuring persuadability. 2 scales did not show sufficient internal consistency and some items showed low item-scale correlations. The 5 successfully developed scales can be used to estimate users susceptibility to certain persuasive strategies. Designers of persuasive technology can identify their intended user groups and use the questionnaire to extract the most effective persuasive strategies to be incorporated in the technology. However, a limitation of this work is that self-report questionnaires are highly susceptible to socially desirable answers and have disadvantages against other methods: Kaptein and Eckles have shown in 2012 [6] that these meta-judgemental measures of personality do a poor job in explaining heterogeneity of responses to persuasive strategies, but can – in addition to demographics and operative measures – still provide additional information. This opens a future challenges in improving these meta-judgemental measures until they can fruitfully complement or even compete other source of data.

Table 1. The persuadability inventory (Final scales and items in black, removed ones in grey)

Rewards (.818) (Cronbachs Alpha)	**ISC**
1. It is important to me that my actions are rewarded.	.593
2. It is important for me to see my success before me.	.394
3. I put more ambition into something, if I know I am going to be rewarded for it	.669
4. I do more work, when I know that I will get something for it (something materialistic).	.673
5. I am willing to change myself if I get rewarded.	.478
6. Rewards motivate me.	.722
Competition (.658)	
7. I push myself hard, when I am in competition with others.	.465
8. I would like to participate in Quiz shows, where I need to assert myself against other people.	.381
9. Generally I am more ambitious than other people around me.	.407
10. I am afraid to be seen as a loser.	.109
11. It is important to me to be better than other people.	.592
12. I like competitive sports (for example racing).	.476
Social comparison (.758)	
13. It is important to me to be equal in comparison to others	.497
14. I like to compare myself to other people.	.624
15. Before I do something, I want to know how other people have done it, so I can feel more save.	.438
16. It is important to me to know what other people are doing.	.522
17. It is important to me, what other people think of me.	.508
18. I adapt my style to the way my friends dress.	.400
Trustworthiness (.472)	
19. I think carefully about if I trust a system before I use it.	.319
20. I trust information better where the source is specified.	.306
21. I trust the information that I receive from the media. (r)	.058
22. I listen more to a person when I know I can trust her/him.	.185
23. It is important for me to be precisely informed about things that I need to do, before I do them.	.446
24. I follow the advice from people that I believe are trustworthy.	.139
Simulation (.368)	
25. I often imagine how the earth will look like in the future.	.266
26. I often imagine how it would be to look differently.	.088
27. I like it when things are well illustrated, so I can get a better picture of things.	.170
28. I find it interesting to know how things work.	.119
29. It is important to me to see what influence my actions have on my surroundings.	.151
30. I change my behavior more, when the results of that change are well illustrated.	.275
Reduction (.361)	
31. I take a detour when I go shopping, if it helps me save money.	.043
32. When the operation of a device is problematic and complicated, I do not use it. (r)	.115
33. I need clear facts to make a decision.	.170
34. When I see the benefit of an action, I am more willing to perform this action.	.263
35. I'm easily willing to follow an instruction that is clear and simple.	.264
36. I prefer to make my decisions using precise information.	.273
Social Learning (.707)	
37. I often modify myself to other people.	.617
38. I ask for advice from other people, before I make a decision.	.319
39. I adopt my behavior quick to the model of other people.	.581
40. When I don't know something, I rather look on the internet or in books, than rely on advice from other people.	.029
41. I adapt my behavior to other people around me.	.594
42. I take other people as role models for new behaviors.	.534

Acknowledgments. This work was partially funded by the European Union Seventh Framework Programme (FP7) under grant 288466 (PEACOX; http://www.project-peacox.eu) and by the Austrian project "AIR – Advanced Interface Research" funded by the Austrian Research Promotion Agency (FFG), the ZIT Center for Innovation and Technology and the province of Salzburg under contract number 825345.

References

1. Fogg, B.J.: Persuasive Technology - Using Computers to Change What We Think and Do. Morgan Kaufmann (2002)
2. Döbelt, S., Schrammel, J., Subasi, Ö., Tscheligi, M.: Applying the Implicit Association Test for the Evaluation of Persuasive Technology. Persuasive Technology - Poster Proceedings (2010)
3. Kaptein, M., Markopoulos, P., de Ruyter, B., Aarts, E.: Can You Be Persuaded? Individual Differences in Susceptibility to Persuasion. In: Gross, T., Gulliksen, J., Kotzé, P., Oestreicher, L., Palanque, P., Prates, R.O., Winckler, M. (eds.) INTERACT 2009. LNCS, vol. 5726, pp. 115–118. Springer, Heidelberg (2009)
4. Kaptein, M.C.: Personalized persuasion in Ambient Intelligence. Journal of Ambient Intelligence and Smart Environments (2012)
5. Kaptein, M., De Ruyter, B., Markopoulos, P., Aarts, E.: Adaptive Persuasive Systems: A Study of Tailored Persuasive Text Messages to Reduce Snacking. ACM Transactions on Interactive Intelligent Systems (2012)
6. Kaptein, M., Eckles, D.: Heterogeneity in the Effects of Online Persuasion. Journal of Interactive Marketing 26 (2012)
7. Torning, K., Oinas-Kukkonen, H.: Persuasive System Design: State of the Art and Future Directions. Persuasive Technology (2009)
8. Cialdini, R.B.: Harnessing the Science of Persuasion. Harvard Business School Publishing (2001)
9. 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)
10. Bauer, R.: Self-Confidence and Persuadability: One More Time. Journal of Marketing Research 7 (1970)
11. Nahemow, L., Bennett, R.: Conformity, persuasibility and counternormative persuasion. Sociometry 30 (1967)
12. Cacioppo, J., Petty, R., Kao, C.: The efficient assessment of need for cognition. Journal of Personality Assessment 48 (1984)
13. MacKenzie, S.B., Podsakoff, P.M., Podsakoff, N.P.: Construct measurement and validation procedures in MIS and behavioral research: integrating new and existing techniques. MIS Quarterly 35 (2011)
14. Hassenzahl, M., Burmester, M., Koller, F.: AttrakDiff: Ein Fragebogen zur Messung wahrgenommener hedonischer und pragmatischer Qualität. Mensch & Computer, 187–196 (2003)

The Scrabble of Language towards Persuasion: Changing Behaviors in Journalism

Cate Dowd

The University of New England, Armidale, 2350, Australia
cdowd2@une.edu.au

Abstract. The rise of social media influences in the online journalism domain suggests that new learning systems are required to modify the behaviors of journalists. The design of future systems can be explored as game concepts and guided by an emerging ontology for journalism. For these ends this paper identifies vocabulary, concepts and emotions in the domain, and vital intersections with social media, such as crowdsourcing. Data from participatory workshops with journalists is applied to new synthetic player ideas, using Hoare logic. It is also lightly structured for a starting ontology for journalism, to inform how a synthetic player system might persuade a journalist to check their behaviors. It prompts the core values of journalism, such as obtaining opposing views, and prompts critical engagement with crowdsourcing, before declaring a story newsworthy. The system includes contextual emotions, which may vary from inspiration and curiosity to anxiety, due to not having a story.

1 Introduction

Journalists in recent years have adapted to major changes brought about by the rise of digital media, including a rise in participatory culture and complex intersections with social media. Newsgathering, editing, and publishing processes are now carried out in convergent multimedia environments that include various database systems. For example, Semantic Web search engines are now used by broadcasters to improve searching for recent and archived stories. Journalist's also use software when monitoring, filtering and gathering news via social media. The changes in journalism practice led to new policies and guidelines for journalists engaging with social media from 2008 onwards. However, they are not complete. Problems still arise for journalists, for both digital natives [1] and established journalists negotiating new systems and social media with traditional values. Practice is compounded by the speed of real-time communication and mobile devices, which impacts on the verification and development of stories. To help address some of these problems the domain must be better understood. This paper builds a knowledge base of journalism including an emerging ontology for the domain with social media contexts, which in turn has informed new ideas for systems to help check and change behaviors.

The paper starts with unfortunate journalism events in late 2012 that led to the resignation of the director general of the BBC. The events included issues at the

S. Berkovsky and J. Freyne (Eds.): PERSUASIVE 2013, LNCS 7822, pp. 39–50, 2013.

nexus of social media, newsgathering and publishing. Related issues began much earlier, and include concepts like crowdsourcing. In 2008 participatory workshops with journalists and other participants [2] generated data for learning systems, for journalism. The research anticipated that future systems would be Semantic Web applications or intelligent agents, such as synthetic players, deployed in Persuasive or Serious Games to alert journalists about behaviors. A synthetic player idea was then conceived by combining data from the domain with emotion-based inputs from human players. This data would also be encapsulated in an ontology for journalism and be reusable for other positive changes in journalism.

2 The Changing Domain of Journalism

In November 2012 the director general of the BBC, George Entwistle, announced his resignation after less than two months in the position. This event raised various issues on the impact of social media on journalism. The event and some associated social media factors are captured in the following extract by a long-term reporter at the BBC, Robinson:

> Labour former culture secretary Ben Bradshaw, an ex-BBC correspondent, said: "Given that in my view George Entwistle - who was a good man, a decent man trying to do his best, only a few weeks in the job - was grossly, grossly let down by people beneath him, and I don't think supported enough by people in the Trust, I think it cannot end here." On 2 November Newsnight reported abuse victim …claims that a leading 1980s Tory politician was an abuser in north Wales, but he withdrew his accusation a week later, saying he had been mistaken. [The Tory politician], although not named on Newsnight, was identified on the internet as the subject of the allegations. He said the claims were "wholly false and seriously defamatory". Mr Entwistle had faced mounting criticism of his response to the programme in an interview …on BBC Radio 4's Today programme. He was criticised for not knowing about the programme at the centre of the row until after it screened and for not being aware of a newspaper article which revealed the mistaken identity, and for not knowing about a tweet saying Newsnight was poised to broadcast the revelations. [3]

The Newsnight story indicates that journalists need to ask more questions, check facts and offer rights of reply, in order to verify a story, but it also shows there is a lack of policy to protect journalists where information is re-contextualized via social media that ‘infers’ someone or something not actually said by a journalist. Inference via social media is not the responsibility of journalists, when they have not mentioned any particular name in a story. This scenario introduces a concept in journalism that can now be named and added to an ontology for journalism. The concept could be referred to as an *inference risk*. This story highlights other problems across journalism and social media. For example, even though journalists use software such as Twitterfall™ to help monitor trends on micro-blogging sites like Twitter, it is impossible to expect journalists or a director, to know about every tweet. In 2012

Twitter generated '400 million tweets per day' [4]. Interestingly, there was one person at the BBC, and at board management level, who was aware of the tweets linked to the Newsnight story before it went to air, but he did not inform the programmers about the tweets. BBC News reported the situation in the following way:

> Mr Entwistle was criticised for not knowing about the north Wales programme until after it screened, for not being aware of a newspaper article which revealed the mistaken identity, and for not knowing about a tweet saying Newsnight was poised to broadcast the revelations. Lord Patten said he was aware of the tweet, which mentioned Conservative politicians, but said it would have been "grotesque interference" if he had contacted the programme then. [5]

An ontologist may ask, what exactly is 'grotesque interference'? Is it a domain concept, or is it rhetoric? Interference that prevents further litigation against a national broadcaster is surely worthy of analysis. The summary and actions following the events at the BBC were summarized in the Ken MacQuarrie report [6]. The summary included organizational change and temporary approaches to management, such as a return to 'single management to deal with all outputs, [and] in order to address the pressure on the Newsnight team, Karen O'Connor agreed to take on the role of Acting Editor of Newsnight'[6]. However, there was no mention of social media issues that had already impacted on the domain. Shortly following the events a note of caution about using social media at the BBC came from the Acting Head of News, Unsworth, reported in the following:

> Referring to the "tumultuous and very sad events of the past few days", Ms Unsworth sent an email to the corporation's staff saying: "It would be helpful if some of our problems were not played out publically across social media and in the pages of the national press. [7]

The comments by Unsworth confirm the clout of social media and its potential to negatively impact on an organization. In a very short time, the best of journalism was forced into a public relations mode and had to adopt the best in crisis communication methods to survive. The events highlighted the need to understand the universal changes in journalism, brought about by social media, towards the design of learning systems to help modify behaviors in the domain.

2.1 The Rising Influence of Social Media in the Journalism Domain

The early social media influences on mainstream journalism anticipated shifts in work processes and 'competition for jurisdiction over 'work tasks' [8:480]. In some ways this was true, but on the flipside new tasks and occupations were created to meet the rise of social media and new media technologies. Bloggers were also 'dependent on mainstream news media for original reporting, a situation that strengthened the position of mainstream journalism' [8:483]. However, by 2012 'original reporting' in mainstream news, evident in the Newsnight story, was not always reliable and the order of things shifted slightly, changing the certainty that 'the order of things in broadcast is 'filter, then publish' [and] the order in communities is 'publish, then

filter' [9:70]. Journalists were perhaps less aware of social media like behaviors creeping into journalism practice, even though they had developed social media policy and guidelines in 2008, which continue to cover areas such as 'blogging and microblogging in terms of integrity and impartiality for those working in news and current affairs' [10]. These guidelines include levels of moderation, for example a pre-moderation approach [11] applies for participation in the area of children's media.

The BBC was also the first to produce an 'Editors' Blog' and noted that 'despite the fears of some that a new openness would lead to embarrassing disclosures … the experiment has been judged at least worthwhile' [12]. The increased engagements with the public online also brought crowdsourcing directly into the journalism domain. Crowdsourcing began from 'cut prices in photos in the US due to distributed labour' [13]. It was a business process understood in terms of 'a theory of crowd wisdom, … outperforming industry faster and cheaper than even the top minds in the fields' [14]. The term was integrated into the vocabulary of journalists, with newly contextualized meaning evident in the conviction that 'one who finds & interviews a source using Twitter™, or crowd-sources case studies via the medium, is certainly reporting" [15]. However, reporting alone is not journalism, and crowdsourcing does not necessarily include sources of information, which is a good starting point for looking at journalism checks against the values of the domain.

The boundaries of information and sources for stories are not always clear. For example, in 2008 a blog called "I've said too much' [16] became an indirect source of information flowing into twitter™, which then washed up into the BBC as a running account of the bomb attack crisis in Mumbai. At the time journalists regarded these flows of information as unconfirmed reports. A recurring issue for journalists using social media is the capacity to verify a situation, and any confusion created by social media can also impact on a journalist's life, especially in a conflict zone. The BBC also use live event pages, which 'allow readers to keep up with the latest information circulating about a story, giving readers a sense of how people are reacting and what they are saying, in real time' [17]. This also helps journalists make sense of a crisis situation and would be vital in a behavior system that encourages critical analysis. For that reason it is included in the emerging ontology presented later in this paper.

3 Making Sense of the Domain via Simulation and Ontologies

The simulation of journalism in the era of social media is largely unchartered territory. This section introduces ontologies that may inform future learning systems for journalism. It explains the need for identification of domain concepts and related ontology approaches that can inform Persuasive and Serious Game design ideas.

Serious Games aim to represent the 'epistemic frame of the professional' [18] and are sometimes developed to simply help understand situations by exploring a domain. This idea is recognized in various European Infrastructure simulation projects that consisted of 'case studies on physical infrastructure projects' [19]. The simulations helped 'to sharpen…insights and tools before they were applied to real-world problem with stakeholders. [19] and to improve understanding of social issues before implementation. They 'represent the mental maps and problems using models of reality' [19:28] and help to satisfy stakeholders prior to implementation. Simulations

can deviate from 'accurately representing real phenomena' [20], but still explore models of reality.

The design ideas that follow in this paper were informed by participatory workshops with journalists, mentioned earlier (see Introduction). The data was then refined to include situated emotions experienced by journalists. The system ideas help make sense of journalism in the era of social media. Ontology engineering approaches were also used to elicit vocabulary and for naming tacit things in the domain, an approach that helps to make sense of journalism for future systems, and is explained further in the next section.

3.1 Domain Vocabulary and Concepts: Ontologies and the Semantic Web

The root meaning of an ontology is 'derived from [a] Greek word, a form of the verb "to be", that means being or existence [21:10]. In computing an ontology is defined as 'an explicit formal specification of a conceptualization' [22]. This definition combines the philosophical meaning with 'a particular system of categories accounting for a certain vision of the world…[where] the system does not depend on a particular [computing] language'[23]. An ontology approach uses 'specific vocabulary…to describe a certain reality [and] in the simplest case, an ontology describes a hierarchy of concepts'. [23:2]

Before attempting the scrabble of language for Semantic Web applications it is useful to know that digital objects identifiers (DOI) are a form of a URI (Uniform Resource Identifier), and that they have 'faceted' associations i.e. the object has a relationship to other related objects. Abstract associations of objects were first evident in the classification of printed books, explained by Ong in the following:

> …a book was sensed as a kind of object which 'contained' information, scientific, fictional or other, rather than, as earlier, a recorded utterance…Each individual book in a printed edition was physically the same as another, an identical object, as manuscripts were not….with print, two copies of a given work did not merely say the same thing, they were duplicates of one another as objects. [24]

Library classification systems have informed the evolution of frameworks for ontologies, notably via the Resource Description Framework (RDF), which is the base for structuring data in libraries. RDF like statements can be derived from an ontology in the form of axioms to assist reasoning about resources. These logical statements used in the Ontology Web Language (OWL) help to 'formulate, exchange and reason with knowledge about a domain of interest' [25]. They constrain possible meaning as multiple statements are parsed using class assertions. In contemporary journalism social media could become a class represented as: ClassAssertion (a:Social_mediaMember a:twitter), which means Twitter is a member of Social_Media.

RDF statements can include the geographic location of a media resource, which may be described further in terms of current or archived media. The statements have a uniform structure of three parts: subject, predicate and object, and include 'facets

when used in the OWL linked as sources, properties and targets' [26]. The three parts are like empty containers for bound data, referred to as triples, which provides the means for linking data via a controlled vocabulary. For that end the refinement of vocabulary in the journalism domain begins in this paper, steering towards language that is meaningful for future Semantic Web applications.

4 Domain Behaviors, Language and Emotions

A typical business process model would not contain information about the emotions of actors in a process. However, emotions can be exploited to improve engagement in a game. For simplicity reasons the emotions in this study were derived from a core of '3 to 11 primary or basic emotions, which includes fear, anger and sadness' [27]. Although it is not wise to restrict human emotions to 'two-state possibilities' [28], for the purpose of a synthetic player, knowledge of typical emotions before and after professional interactions could be useful. Related data from participatory workshops with journalists [2] were collated and combined by the author into a single table (see Table 1).

Table 1. A list of role-based tasks or actions for journalists with emotions named before and after tasks

Role	Task or Action of Journalist	Before interaction	After interaction
Newsgathering	Ascertain story - make calls	Curiosity	Satisfied
Newsgathering	Find opposing views	Frustration/panic	Completion/relief
Editor	Manage & prioritize news	Stressed	Satisfied
Editor	Read papers monitor news	Mixed	Satisfied
Newsgathering	Attend news conference	Anticipation	Satisfied
Writer	Find out what story is	Anxiety	Inspired
Newsgathering	Record facts, direct crew	Anxious, aggressive	Satisfied
News Editor	Checking for defamation	Anxiety	Relief
Editor	Reporter knowledge – canon	Anxious/excited	Anxious/glad
Editor	Decide approach to story, gather tips	Bargaining, curiosity	Understanding
Camera person	Get exclusive shot	Thrill	Smug
Editor	Apply values to news angle	Pressured	Smug/satisfied
Sub Editor	Headline meeting	Anxiety	Pleasure
Editor	Ask questions – police estimates of crowd numbers	Interested	Satisfied
Camera person	Get story information from journalist	Boredom	Contempt
Editor	Manage & prioritise news	Pulled in different directions	Bruised
Writer	Order information using editing tool	Panic	Satisfaction
Writer	Ascertain story, establish opinions/length & check facts	Anticipation, curiosity	Understanding, clarity

Table 1. (*continued*)

Photographer	Ascertain story, framing pics, options	Clarity, curiosity	Satisfaction, completion
Section editor	Establish story options, decide final story & layout	Bargaining, curiosity	Understanding, clarity
Editor	Rank stories for news bulletin	Chaos	Organization
Correspondent	Persuade Exec Producer to do story you believe is good	Mixed	Delight
Newsgathering	Ascertain story - read press releases and wire service	Curiosity	Appreciation
Correspondent	Break new story - get a document no one else has	Competitive	Smug
Newsgathering	Decide approach - talk to editor chief	Bargaining	Understanding
Correspondent	Talk to source(s), check facts, negotiate with editor	Curiosity, bargaining	Clarity
Newsgathering	Truth communicating with the subject	Anticipation	Inspired
News Editor	Monitor story from an agency	Satisfied	Excitement
Sub Editor	Legal advice: send script to in-house adviser via email	Hurried/seeking	
Editor	Check accuracy of story by asking reporter source of information		

The data from Table 1 was refined further for ideas for synthetic players. The contextualized data could be used to develop a database of emotion representations, bound for machine-readable purposes, and linked to objects via logic systems. A base approach for the application of logic is introduced in the next section.

4.1 The Construction of Artificial Behaviors and Emotions Using Logic, for Learning about Journalism

Algorithms that explore emotion states as potential matches between a pre-condition and a post-condition appear to be well suited to Hoare logic [29], and could be applied for game systems. Hoare logic is a formula for working with data in computing that consists of a pattern of triples, represented in the following way:

$$\{P\}\ C\ \{Q\} \tag{1}$$

The symbols of Hoare Logic are understood as: 'P and Q are assertions and C is a command. P is the pre-condition and Q is the post-condition. P and Q are formulas in predicate logic: Whenever P holds for the state before execution of C, then Q will hold afterward' [29]. If this logic were applied to a synthetic player engaging in journalism newsgathering with displays of emotion states as extracted from the collated data shown in Table 1, and in context of ascertaining a story, it could consist of a triple that looks like this:

$$\{\text{Curiosity}\}\ \text{publish story}\ \{\text{satisfied}\} \tag{2}$$

If this triple was applied to a synthetic player the logic may suggest that the synthetic player always shows satisfaction once a story is published. However, an actual player would need additional challenges. Another algorithm could be generated that requests from the player an exclusive story! The representations to match the conditions could include the following triples:

{Competitive} scoop please {smug} (3)

Each level of an algorithm could prompt a player towards a higher level of understanding of the journalism domain. These examples use just one form of logic towards 'intelligent behavior [that] can be achieved through processing symbolic structures representing increments of knowledge [30:33]. They are starting statements towards 'formal symbols to represent a collection of propositions [31:4], which could hold true for the journalism domain.

Low-level intelligent synthetic players for the journalism domain could be built up from many proposition statements and various applications of logic, as introduced in this section, and by using the bound behaviors and emotions in Table 1.

The following section presents a concept created by the author derived from data about journalism processes, which may be useful for sub-system ideas for a future Persuasive or Serious Game, or for better understanding of contemporary journalism.

4.2 Emotions for Driving Finite State Machines: Robo-journo

The context-based emotions (see Table 1) and other data are synthesized in this section into a model of typical interactions. The model includes finite states and transitions between states, which may be used for synthetic players or mobile participants. Decision making associated with a Finite State Machine (FSM) may be used at 'strategic, tactical or operational' [32:116] levels in a game. Such systems may include pattern recognition to assist 'strategic decisions made for a long period of time [and] based on a large amount of data' [32:117]. If transition inputs could be partly based on player emotions, it might be possible to validate further typical emotions in the journalism domain, from inputs over time, using pattern recognition.

The conceptual design of a synthetic player such as Robo-journo (see figure 1), brings to the foreground the flow, logic and emotions associated with journalism processes, including social media elements like crowdsourcing. Robo-journo manages press releases in a game space dedicated to newsgathering and the system is partly driven by typical emotions found in the journalism domain.

Hypothetically, a press release processed by Robo-journo could have been generated via email or arrived as a promotion kit. Either way Robo-journo enters a cycle of typical, but limited states, associated with the life of the press release, and in context of journalism tasks and associated checks via social media, towards verification of a story. Once the process is complete Robo-journo returns to base and generates more press releases.

The Robo-journo machine firstly determines if a role-playing journalist is in a good state to play, and if so, it moves to the next state where the press release prompts a story. The player could show curiosity about the potential of the press release for

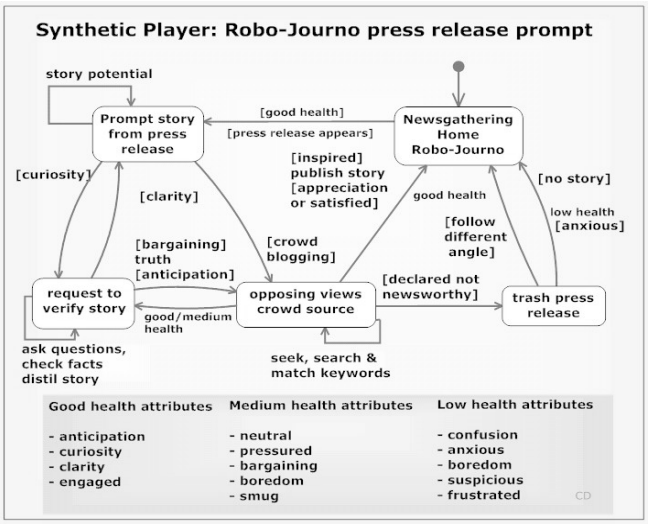


Fig. 1. A Finite State Machine schema for a synthetic player which collaborates to modify behaviors of journalists by responding to emotion inputs

creating a news story. If the journalist displays curiosity then Robo-journo requests that the player verify the story. To move to the next state, a role-playing journalist must bargain or negotiate to get the story they want. If there is a lack of clarity about a potential story, then Robo-journo may request opposing views. The synthetic player could also start a process of crowd sourcing using SEO techniques and tools. If the journalist is inspired and writes a story, then Robo-journo returns to base and starts the process over with another press release.

Robo-journo is not interested in ethics. If a journalist is not inspired to write a story and declares a press release is not newsworthy then it puts the press release in the trash and returns home. A role-playing journalist may use the press release to take takes another angle for a story. If there is no story from the press release a journalist may indicate they have wasted time, show frustration, and the synthetic player could be given a low health status.

Robo-journo knows certain things about journalism processes and responds to emotions to drive the player and system. The concept could be refined bearing in mind that 'the scope of an FSM be limited so that software modularity is more adaptable [32:134-135]. Robo-journo should be able to persuade journalists to change behaviors, using additional data that could be derived from a domain ontology.

5 An Ontology for Journalism: Towards the Semantic Web

The emerging ontology for journalism in this section includes elements referred to as *things*. A social media 'thing' has relationships to other things in the journalism domain. There are 16 classes in the journalism ontology (figure 2) and several classes

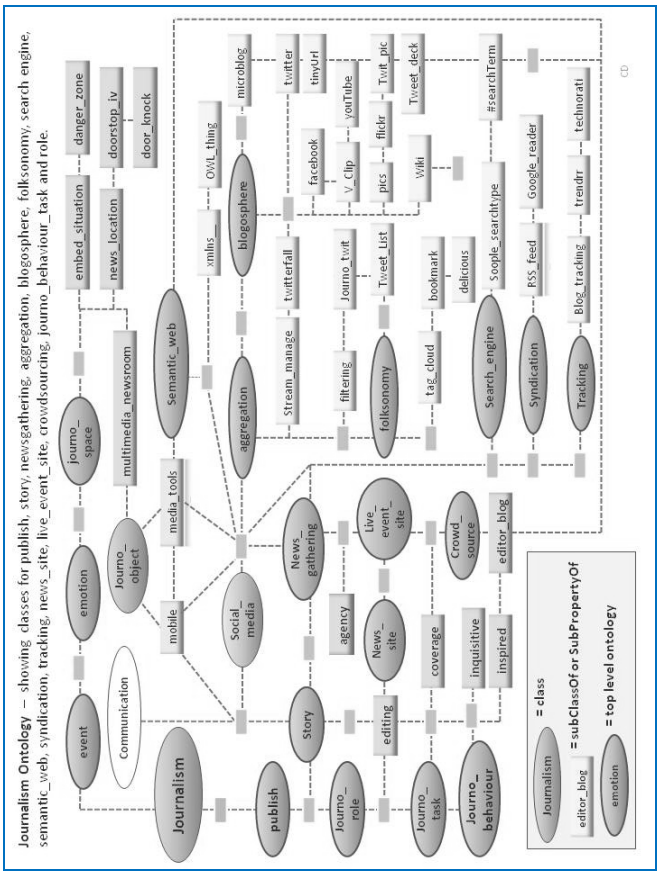


Fig. 2. A emerging ontology for journalism with social media intersections

not shown, such as: journalism_guideline and source_information. There are also 4 top-level classes: event, emotion, journo_space and journo_objects.

The journalism ontology has various slots for sub-class properties, such as character attributes from the journo_behaviour class. A vital feature of online journalism is the live_event_site and this has a hierarchical relationship to newsgathering and crowd sourcing. Aggregation stems from a larger class of social media, and is in close proximity to XML [33]. Semantic web tagging systems, such as DBpedia [34] which enables the use of a 'tag displayed as a link to an aggregation page for [a] concept' [35] are at the top level, as Semantic searching via news sites is now commonly used by journalists and media consumers.

OWL may not be ideal for developing future systems, but the ontology approach forces concrete descriptions. A snippet of code (figure 3) using OWL 2 syntax indicates how data could begin to be coded with individual owl members associated with newsgathering, transforming ontology data to code.

```
1. <owl:Class rdf:about=
http://journalism.ont/story/newsgathering/socialmedia/crowdsou
rce”>
2. <owl:Class rdf:ID=“newsgathering“>
3. <rdfs:subClassOf rdf:resource="#crowdsource" />
4. <owlx:Individual owlx:name=“twitterfall“>
5. <owlx:Individual owlx:name=“#searchterm“>
```

Fig. 3. A snippet of OWL code by with terms extracted from an emerging ontology for journalism to begin reasoning about resources

For the purpose of future learning systems for journalism, the ontology approach has a long way to go, but the structuring of domain data can lead to new possibilities for changing behaviors, using rich data approaches.

6 Conclusion

This paper has captured several major transitions in the journalism domain and highlights social media impacts on journalism. These impacts suggest the need for persuasive learning systems for journalism. Participatory workshops with journalists helped to make sense of the domain for future Semantic Web and game applications and this paper has combined emotions with logic for synthetic players to help check the behaviors of journalists in the era of social media.

References

1. Prensky, M.: Digital Natives, Digital Immigrants. On the Horizon 9(5), 1–6 (2001)
2. Dowd, C.: The Scrabble of Language for Real-time Crisis Communication: Serious Game Design for Journalism, Social Media and Public Relations. Unpublished PhD thesis, CSU (2011)
3. Robinson, N.: http://www.bbc.co.uk/news/uk-politics-20287066
4. CNet, http://news.cnet.com/8301-1023_3-57448388-93/twitter-hits-400-million-tweets-per-day-mostly-mobile/
5. BBC, http://www.bbc.co.uk/news/uk-20286198
6. BBC, Ken MacQuarrie report: Summary of findings, http://www.bbc.co.uk/news/uk-20306096
7. Raynor, G.: http://www.telegraph.co.uk/culture/tvandradio/bbc/9674996/BBC-news-chief-pleads-with-staff-not-to-Tweet-about-our-problems.html
8. Lowrey, W.: Mapping the journalism-blogging relationship. Journalism 7, 477–500 (2006)
9. Lasica, J.: http://www.nieman.harvard.edu/reports/contents.html
10. BBC, http://www.bbc.co.uk/editorialguidelines/page/guidance-blogs-personal-full
11. BBC, http://news.bbc.co.uk/2/hi/help/4180404.stm
12. Barron, P.: http://www.bbc.co.uk/blogs/theeditors/2007/06/editors_blog_the_first_year.html

13. Howe, J.: http://www.wired.com/wired/archive/14.06/crowds.html
14. Brabham, D.: Crowdsourcing as a Model for Problem Solving: An Introduction and Cases. Convergence: The International Journal of Research into New Media Technologies 14(1), 75–90 (2008)
15. Posetti, J.: http://www.j-scribe.com/2009/11/twitter-as-journalistic-tool-drilling.html
16. Shepherd, L.: I've said too much blog, http://www.lllj.net/blog/?page_id=283 (no longer available on-line)
17. Cellan-Jones, R.: http://www.bbc.co.uk/blogs/technology/2008/12/twitter_the_mumbai_myths.html
18. Shaffer, D. W.: http://epistemicgames.org/eg/?p=414#more-414
19. Mayer, I., Veeneman, W.: Games in a World of Infrastructures: Simulation Games for research, learning and intervention. Eburon, Delft (2002)
20. Crawford, C.: What is a Game? Washington State University, Washington, http://library.vancouver.wsu.edu/art-computer-game-design
21. Luger, G.: Artificial Intelligence Structures and Strategies for Complex Problem Solving. Pearson Education, Boston (2009)
22. Gruber, T.: A Translation Approach to Portable Ontology Specifications. Knowledge Acquisition 5(2), 199–220 (1993)
23. Guarino, N.: Formal Ontology and Information Systems. In: The FOIS 1998, Trento (1998)
24. Ong, W.: Orality and Literacy. Routledge, London (1982)
25. W3C, http://www.w3.org/TR/2009/REC-owl2-primer-20091027/#Classes.2C_Properties.2C_and_Individuals_.E2.80.93_And_Basic_Modeling_With_Them
26. W3C, http://www.w3.org/TR/2009/REC-owl2-rdf-based-semantics-20091027/#Vocabulary_Terms
27. Plutchik, R.: The Nature of Emotions: Human emotions have deep roots, a fact that may explain their complexity and provide tools for clinical practise. American Scientist 89(4), 344–350 (2001)
28. Adams, E., Rollings, A.: Game Design and Development. Pearson Prentice Hall, Upper Saddle River (2007)
29. Hoare, C.: An Axiomatic Basis for Computer Programming. Communications of the ACM 12(10), 576–583 (1969)
30. Alesso, P., Smith, C.F.: Thinking on the Web: Berners-Lee, Gödel, and Turing. John Wiley & Sons Inc., New Jersey (2006)
31. Brachman, R.J., Levesque, H.J.: Knowledge Representation and Reasoning. Elsevier, Amsterdam (2004)
32. Smed, J., Hakonen, H.: Algorithms and Networking for Computer Games. John Wiley & Sons, England (2006)
33. W3C, http://www.w3.org/XML/
34. DBpedia, http://dbpedia.org/About
35. Raimond, Y., Scott, T., Sinclair, P., Miller, L., Betts, S., McNamara, F., et al.: http://www.w3.org/2001/sw/sweo/public/UseCases/BBC/

A Customisable Dashboard Display for Environmental Performance Visualisations

Daniel Filonik[1,2], Richard Medland[2], Marcus Foth[2], and Markus Rittenbruch[2]

[1] University of Munich, Media Informatics Group
Amalienstr. 17, 80333 Munich, Germany
daniel.filonik@stud.ifi.lmu.de
[2] Urban Informatics Research Lab., Queensland University of Technology
130 Victoria Park Road, Kelvin Grove QLD 4059, Australia
{richard.medland,m.foth,m.rittenbruch}@qut.edu.au

Abstract. We conducted an exploratory study of a mobile energy monitoring tool: The Dashboard. Our point of departure from prior work was the emphasis of end-user customisation and social sharing. Applying extensive feedback, we deployed the Dashboard in real-world conditions to socially linked research participants for a period of five weeks. Participants were encouraged to devise, construct, place, and view various data feeds. The aim of our study was to test the assumption that participants, having control over their Dashboard configuration, would engage, and remain engaged, with their energy feedback throughout the trial.

Our research points to a set of design issues surrounding the adoption and continued use of such tools. A novel finding of our study is the impact of social links between participants and their continued engagement with the Dashboard. Our results also illustrate the emergence of energy-voyeurism, a form of social energy monitoring by peers.

Keywords: energy monitoring, environmental sustainability, persuasive technology, domestic environments, households, urban informatics.

1 Introduction

This research project investigates the potential of real-time data stream composition and visualisation in the context of energy monitoring. The ultimate goal is to provide end-users with tools that enable them to combine information from various data sources – such as domestic sensors, social media, or public transport – in order to support day-to-day activities and allow them to make informed decisions. As a first step towards exploring this concept, a prototype of a customisable Dashboard system was developed and evaluated in a user study.

The widespread deployment of hardware for sensing various types of data is becoming technologically and economically feasible. As a result, citizens are increasingly producers as well as consumers of real-time data streams – a trend that Paulos et al. term 'citizen science' [7]. The purpose of the Dashboard prototype is to give the user a comprehensive and intuitively accessible overview

S. Berkovsky and J. Freyne (Eds.): PERSUASIVE 2013, LNCS 7822, pp. 51–62, 2013.

of data that is relevant to their local household. The study acknowledges that individual users have different preferences with regards to the data that they want to see displayed and the style in which they want it to be presented.

A key contribution of this study lies in the evaluation of the developed Dashboard prototype with actual users, observing their interactions with the system in the environment where they would normally use it. The goal of the study was to determine the persuasive capacity of the Dashboard to allow for different data stream compositions and visualisations. A central question was how participants would make use of customisation aspects of the composable information display and whether it would influence their interactions with the system and their consumption patterns.

2 Related Work

The basic premise of using *Information Visualisation* to display data related to resource consumption is that consumers are lacking information - particularly when it comes to intangible commodities such as electricity. Making consumption data available in an intuitive manner in real-time allows users to better understand the impact of conservation efforts and make informed decisions. Feedback helps to expose consumption which was previously invisible, which is especially the case for domestic energy use. Most commercially available feedback displays rely on an intrinsic rational-economic model, which assumes that people will be motivated by the prospect of saving money [4,12].

In order for feedback to be effective, it is important to carefully consider what to present and how. Studies have demonstrated that it is easier to persuade users by addressing specific behaviours rather than general ones. Therefore, actionable feedback is required that highlights the necessary steps to reach a desired goal. In the case of energy conservation, this requires systems that are capable of analysing consumption and pointing out potential savings [5,8,10].

However, researchers increasingly recognise that the mere availability of information is not sufficient to affect sustained behaviour changes. Nisi et al. observe that many users lose interest in monitoring their energy consumption within weeks and even with the introduction of new visualisations "the novelty effect only [lasts] for less than a week" [6]. Financial incentives can be effective, but many users relapse into old behaviours if the rewards are small in comparison to income and the novelty wears off. A common criticism is that the rational-economic model is not well suited to accurately represent how and why people consume. Strengers illustrates the issue as follows:

> "It is unlikely that most of us, on rising from our slumbers each morning, approach every task 'rationally' by consciously weighing up the costs and benefits of a shower, or ensuring we undertake the most efficient load of laundry." [11]

Strengers elaborates that everyday activities are much more guided by social norms, cultural dynamics, institutional rules, and technological means. If the

consumer does not already hold a conviction to conserve energy, then feedback only informs, motivating neither attitudinal change nor action [3,12,11].

In order to change behaviour, *Persuasive Technologies* have to consider personal attitudes and interests. Extrinsic forms of motivation can act as triggers to foster intrinsic motivations. Further, it is possible to differentiate between several "stages of readiness, willingness and ableness to change" [3]. He et al. draw upon the stages of behaviour change identified as part of the *Transtheoretical Model* to develop targeted motivational strategies. Because users have different motivations – and these motivations change over time – the authors assert that it is not feasible to develop a 'one-size-fits-all' solution [3,12].

In their evaluation of *StepGreen.org* – a website designed to encourage energy saving behaviours – Mankoff et al. acknowledge that "no one visualization fits all users and contexts" [4]. The authors plan to address these varying needs by developing a more flexible and adaptive solution. A common approach for achieving this goal is to support end-user customisations, allowing users to personalise their interface. Interaction with the application invokes self-reflection and elicits a sense of freedom, which can increase intrinsic motivation. Further, a personalised interface is a more effective motivator than one that displays general information [3,4,13].

There is great potential for *Information Visualisation* to engage people with their resource consumption when it acts as the basis for social interactions. The *Wattsup* application allows users to visualise and share energy consumption data on Facebook. In a study of the application, Foster et al. observed that a socially enabled version was "more enjoyable and more effective than individual monitoring" [2]. The users engaged in banter and competition, which proved to be motivating. A study by Vande Moere et al. that exposed energy consumption on house façades also echoes the beneficial effects of making personal information publicly available, such as peer pressure and healthy competition [2,12].

This existing research reaffirmed our initial assumptions about the potential value of customisation and social interaction for *Persuasive Technologies*. It encouraged us to specifically explore three concepts:

Customisable Data Sources. Do users appreciate the ability to select data sources that are relevant to their personal interests?

Customisable Visualisations. Do users gain a greater understanding of the data if they can tailor the visualisation to their needs?

Social Sharing. Does the exchange of environmental performance data improve understanding and motivate pro-environmental behaviour?

3 Methodology

The main goal of this study was to develop a Dashboard prototype that could be deployed to users in the real world. Ultimately, the prototype should assist users with their real-world tasks, allowing them to make informed decisions. To maximise adoption, the solution needed to be accessible and user friendly. Therefore, the project was guided by a user-centred design approach.

In accordance with the core principles of user-centred design, the development consisted of several iterations and the results were repeatedly verified with experts and prospective users. The project made use of tools and techniques for rapid development in order to create a succession of incremental prototypes. The main iterations of the development process are summarised in figure 1. A key goal was to employ adequate evaluation methods at each stage of the project. The project culminated in an exploratory user study, which served to verify as-

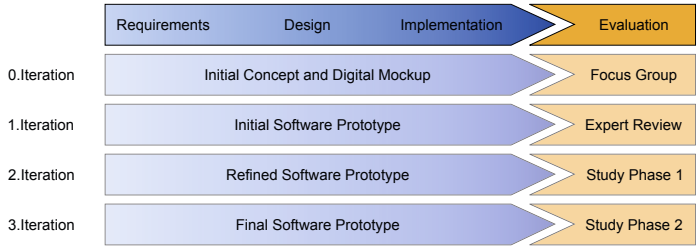


Fig. 1. Iterations during the development

sumptions about the role of the Dashboard prototype in a real-world setting. It was carried out as a field trial, spanning two separate phases. Due to the unique characteristics of the prototype, a holistic evaluation of the user experience took place over time in a natural environment rather than a controlled laboratory setting. The study relied on a mix of observational and quasi-experimental methods to estimate the role of the Dashboard prototype with regards to energy monitoring. Quantitative data was collected in the form of empirical usage statistics, tracking various user interactions with the Dashboard system. However, due to the quasi-experimental design, it was essential to supplement this data with qualitative feedback from participants.

4 Design Process

This section describes the activities that informed the major design decisions and shaped the requirements of the Dashboard system. The main concepts of the prototype were fleshed out during the early iterations. In particular, the focus group and the expert review delivered valuable input for the overall design. In the later phases running up to the user study, the development focus shifted towards refining the design, extending the system and addressing usability issues.

4.1 Focus Group

The focus group comprised a total of 8 participants, 2 female and 6 male, aged between 25 and 43 years. They were recruited from the community surrounding the Urban Informatics Research Lab. Some participants were taking part in the

overarching research program on domestic energy conservation, meaning that they already had energy monitors installed at home. Furthermore, the group included PhD students in *Human-Computer Interaction (HCI)*. Therefore, the participants were mostly computer savvy and already had experience with energy monitoring. This was deemed to be a favourable constellation, because it was possible to draw from their experience and discuss advanced topics, such as aspects of the user interface. They were also indicative of the early adopters of tools such as the Dashboard prototype, since these individuals are likely to be technically adept and sensitised to the issue of energy conservation.

Over the course of the session, participants worked on different tasks, either collectively or individually. Together, participants collected an extensive selection of different kinds of information that were relevant to them on a regular basis. The broad scope and variety of ideas confirmed our initial concept, because users clearly had a wide range of differing information needs.

Later, individual sketches also demonstrated the great diversity in user expectations. Many of them were quite detailed, featuring innovative ideas for combining and visualising information. The participants often made use of the limited space by displaying different kinds of information within a single visualisation. An example of this was a map that displays traffic congestion and overlays the current weather in the style of a weather map. Energy consumption was commonly presented with a line chart displaying data over a fixed period of time, including one notable exception which adopted a glowing orb metaphor.

4.2 Expert Review

Once an initial working prototype was ready, an expert review was conducted to discover potential problems and ensure that the project was heading in the right direction. A new group of 8 experts was recruited for the evaluation, with about half of them encountering the project for the first time. All of the attendees were colleagues from the Urban Informatics Research Lab. Therefore, the group was mostly comprised of PhD students and academics with experience in HCI. The review took on the shape of a formal usability inspection, where "experts hold a courtroom-style meeting [...] to present the interface and to discuss its merits and weaknesses" [9].

Even though the system was still at an early stage, the core components of the system architecture – the web service and the native tablet application – were already in place, providing all of the basic functionality. The demonstration covered the complete system, including the workflows for authorising services and creating data streams. However, the subsequent usability evaluation was primarily focused on the native tablet application. For this purpose, the experts received a tablet device which had the prototype installed and running, allowing them to configure and customise the Dashboard display.

A number of minor usability flaws were identified and addressed during the next round of development. In general, the experts praised the visualisations that were enabled by the system as well as the implemented sharing functionality. They saw great promise in using the system for social comparison of energy

consumption data. The potential of end-user customisation was controversial. In particular, the benefit of combining unrelated information was questioned. Some experts were sceptical whether the presence of certain information, like social media, would motivate users to view the Dashboard display. One participant noted: "If I would like to check my tweets I probably would use the Twitter app rather than a new one." On the other hand, some experts were optimistic about the idea. Another participant pointed out: "There are some feeds that I know I should be checking in a kind of peripheral vision and then there are some things that I know I check all the time." He concluded that, "by combining the things I do all the time with the ones that I know I should, but I never do, you kind of get [the best of both]." As a result, the discussion brought attention to the fact that the concept may appeal to some, but not everyone.

5 System Functionality

This section serves as an introduction to the two main components of the Dashboard system developed as part of this research: the web service and the native tablet application. This system architecture is the result of a deliberate design decision to maintain a clear model-view separation. On the one hand, the web service aims to provide a universal data stream brokerage platform. It allows users to manage collections of relevant data streams from heterogeneous services and exposes them through a unified API. On the other hand, the tablet application acts as a client of the web service, using it to discover and query the data streams that drive its visualisations. Its main purpose is to allow users to compose and visualise data.

5.1 Web Service

As a user, it makes sense to start exploring the Dashboard system by visiting the web service and configuring an initial set of data streams. Informally, a stream can be described as a sequence of items. Each item contains multiple attributes that form a self-contained unit of information, such as a reading from an energy monitor or a message on a social network. This highly generic definition allows to model dynamic content from numerous online services.

The core functionality is accessible from an overview page, which is displayed once a user enters the private section of the web service (see figure 2). This page lists all supported services and the associated data sources. The current implementation includes plugins that integrate data from sensors (*Pachube*), government agencies (*Bureau of Meteorology*), social media (*Facebook*, *Foursquare*, *Twitter*), and generic services (*Google*, *RSS*). Users populate their personal collections with streams from data sources relevant to their interests.

Furthermore, the Dashboard web service is a natural place for providing social features. For this purpose, the service incorporates a basic friend system that allows users to interact with trusted individuals. Since data streams are viewed as resources belonging to individual users, their owners can share access to them with their friends.

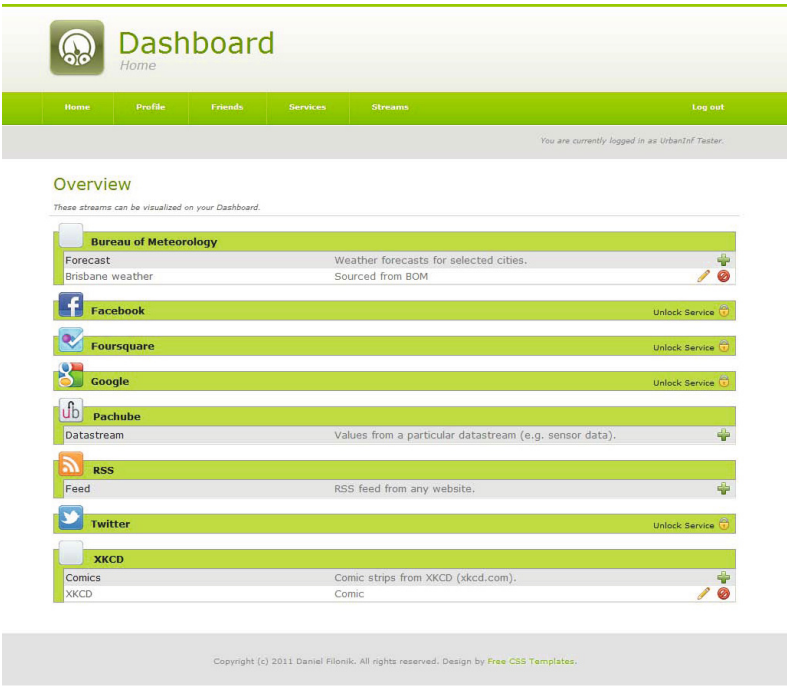


Fig. 2. Overview page showing services and streams

5.2 Native Application

After configuring data streams on the web service, it makes sense to move on to the native tablet application. The familiar concept of widgets is adopted for representing individual visualisation components. Initially, the display is empty, providing users with a blank canvas on which they can arrange their ideal selection of widgets. This supports the key objective of allowing users to tailor the display to their personal needs and interests. Users can choose from a library of widgets that is bundled with the application, which includes charts (*Line Charts*, *Sparklines*, *Stacked Columns*), maps (*Markers*), images (*Slideshows*), as well as textual (*Notes*, *Messages*, *Articles*) and iconic (*Weather*) representations.

Upon authentication, users are presented with their personal Dashboard display (see figure 3). The 'view'-mode is optimised for information retrieval. By design, the display area is limited to a single screen, showing all personally relevant information at a glance. In 'view'-mode, the widgets do not respond to touches, preventing accidental modification. The only possible action is focusing on a certain widget by performing a double tap gesture. In this state, widgets can receive touch input, allowing for interactive visualisations. In order modify their Dashboard, users enter 'edit'-mode with a long press gesture. Visual cues highlight the transition to a different mode of operation where it is possible to configure, scale, and arrange widgets.

Fig. 3. Main screen showing widgets in 'view'-mode

6 Evaluation

During March 2012 a refined prototype was deployed to five new subject-matter experts for a period of three weeks. Participants ranged in age from 27 to 43 with differing technical backgrounds, including HCI and interaction design. Three of the participants were regular *iPad* users. Taking the prototype out of the laboratory and into the homes of participants was the natural next step following our on-site expert review. After the initial setup of the Dashboard with the researcher present providing instructions, the expert users were asked to customise their Dashboard display to suit their needs. During this deployment the directive was to make use of the Dashboard and to identify any issues. Ad hoc feedback was gathered via email correspondence. Structured feedback was gathered using interviews, with each participant interviewed prior to using the Dashboard, and at the end of the three week trial. Using these methods, participants were able to describe in detail the aspects of their interactions with the device, their observations of usage patterns and adjustments to their routines. Participants critiqued the usability of the prototype in post-interviews, and also reflected upon their usage patterns and overall impression of the Dashboard.

In May 2012 a larger five week study was undertaken, deploying the final prototype to members of the local community. The most serious usability issues were addressed to make the system more intuitive for participants. The widget selection was increased and existing widgets were improved using participant feedback. Overall, 13 participants from 12 different homes were sourced as

part of a larger program of research into energy conservation, each having their own *iPad*. These participants already had a *CurrentCost Envi-R* energy monitor installed in their homes, and basic demographic data had been recorded. Again, participants were individually interviewed twice, at the beginning and conclusion of study phase 2. Numerous individual interview questions teased out the function of the Dashboard for each participant. Initial questions sought high face validity, e.g., "what impact, if any, do you think the Dashboard had." Later questions targeted the impact of customisable information visualisation and end user customisation. Additionally, at the conclusion of this deployment, the most socially linked participants were included in a two-hour phase 2 focus group. This second focus group further investigated the persuasive elements of the social sharing components of the Dashboard.

The usability evaluation clearly showed that the prototype, while providing a level of feedback that the users found novel, needed more refinement. The issues revolved around configuration, interface presentation, and requisite diversity of widgets to satisfy their individual requirements. These shortcomings will be addressed in future iterations.

6.1 Social Ties and Voyeurism

In the phase 2 focus group, it became clear that while participants showed preference for different Dashboard widgets, the overriding element that helped to persuade some to remain engaged was the ability to share and compare energy consumption. This theme featured prominently throughout the phase 2 focus group and is exemplified by the following statement by Ronaldo (names have been replaced throughout), "I did notice an uptake in using the Dashboard once I had Fernando and Paolo on there as well. Once I was able to overlay my consumption with theirs, because thats the only app where I can do that, I kind of checked on that more often."

Another theme that emerged was the feeling of *voyeurism* that Fernando commented upon, jokingly stating that he felt the desire to "obscure" his energy consumption, knowing that Ronaldo was watching. Fernando then commented on Paolo's energy consumption stating "Paolo you can tell when [housemate 1 and 2] get home, at 6pm onwards, there's a spike." Zinedine stated "maybe you could see OK, there's a party going on in Fernando's house or something and then follow up." This theme supports previous research publicly exposing energy consumption [12], and offers a persuasive method for engaging participants in the future. The underlying message here was one of comparison with those participants who shared social ties [1], e.g., Ronaldo benchmarking his already low energy use with Fernando and Paolo. The ability to compare or play energy voyeur represent real avenues for motivating energy conservation.

Speaking on this topic Ronaldo related his own experience; the emergence of an informal group of three friends involved in curbing energy use when he became interested in lowering his energy consumption. The group achieved excellent results with all members lowering their energy consumption, eventually installing solar panels and hot water, trading knowledge and appliance usage

patterns through collaborative documents. They reached a maintenance mode after a period of months [3], where their collective learning plateaued and members, having made lifestyle concessions to consume less energy felt that they had reached their goals. The group established a new norm for their energy consumption and behaviours through a learning process facilitated by engaging with socially linked individuals. Ronaldo's knowledge sharing conducted in the phase 2 focus group was enabled by the Dashboard energy widget. The ability to compare real-time energy consumption data with friends over a period of time is a clear pathway for other persuasive tools [2,11].

6.2 Data Stream Composition

The notion of having a customisable Dashboard, and view multiple streams of information in one interface was regarded positively by the participants. Each participant adapted their Dashboard to their own interests, with Fernando displaying the densest Dashboard with RSS feeds (for daily deals and discounts), social media, local news, weather, and multiple energy widgets (both for consumption and photovoltaic production). The aspect of combining widgets to deduce useful information was exemplified by participants with solar panels. Conscious of the return on their sustainable energy investment, these participants re-purposed the weather widget to both predict and deduce the reasons for differing levels of energy production displayed by the energy widget.

Fernando and the other participants agreed that they were seeing their energy consumption frequently by using the Dashboard, but were often just monitoring the pattern displayed on the energy graph. The experience of the participants meant that they were often displaying a maintenance behaviour, noting that "nothing was wrong" and then moving on. Echoing this comment, another participant mentioned in their final interview that a simple red light/green light would be helpful for providing feedback on "excessive" energy use. Future widgets should cater to this information seeking behaviour providing simple, at times ambient visual cues to facilitate maintenance behaviours for participants.

Participants viewed displaying multiple configurable information sources on a single display as useful, though in certain cases they were unable to articulate how they would use such a tool on a day-to-day basis. This may represent a source of bias in their responses or familiarity with the interviewer. As this study is positioned as exploratory and experimental the day-to-day use of the Dashboard was not applied as a metric of success or failure.

6.3 Intrinsic Motivation

When asked what their intent was when viewing the Dashboard, most participants stated that they went to look at the energy consumption widgets. This behaviour was especially prevalent on those with solar panels, which appear to be an indicator of energy consumption awareness. The desire to derive maximum value from previously unmonitored equipment was often the reasoning offered by participants. The budget widget with a configurable total was Zinedine's

favourite. The widget provides a column that fills as a target is reached, and was added based upon feedback in phase one on goal setting. Zinedine had put it to use as a way of tracking solar production commenting, "I'm trying to generate 10 kWh per day, so it just gives me a percentage of how much I've done". The line of thought extended with Zinedine introducing the concept of cross-widget scripting, where if the budget target was reached the power might be shut off for his home or the lights might dim at 50%, acting as an alarm. This presents an interesting next step for this research considering widget mash-ups designed by users to suit their desires. In line with this suggestion Zinedine stated that he would like to be seen as an author of content when another friend adopted one of his shared widgets on their Dashboard. This form of recognition was ideal for Zinedine as long as he was able to see how many people were using it. Applying this notion of participants deriving status from the creation and use of customised widgets by others is another avenue for future research.

Participants were not readily able to assess a change in energy consumption after having used the Dashboard. This was often related to lifestyle norms [11], or for more experienced participants a sense that they were already informed about their energy consumption and had minimised it. For experienced participants the energy widgets did not represent value until the social comparison and budget features were provided. As was described previously, these helped experienced participants remain engaged, perform maintenance behaviours, benchmark their consumption, and caused serendipitous social interaction.

7 Conclusion

Through our exploratory study, we gained a better understanding of the opportunities and challenges of deploying a customisable Dashboard display in a domestic environment. We observed strong indicators that some of our assumptions materialised and provided real benefits for energy monitoring. With regard to customisation, most users appreciated the ability to select visualisations and many of them desired a greater widget selection with deeper data analysis functionality. In particular, users were enthusiastic about compositions of widgets that allowed the to investigate relationships between the data, such as the correlation between weather and energy consumption.

The most promising dynamics with regard to energy conservation arose in the socially enabled condition. Participants that shared their energy consumption with others were likely to exchange expertise and troubleshoot problems. Furthermore, we observed the emergence of energy-voyeurism, a form of social energy monitoring where users – driven by their own curiosity – compared and analysed the energy consumption of others.

In concluding this work, it should be noted that of recurring interest to this work is the need for more research into the inherent interest of individuals in their energy consumption behaviour. A clear avenue for future research is in developing a toolset for assessing the likelihood for an individual of adopting or rejecting energy conserving behaviours. This assessment should also include the relative perceived impact upon lifestyle.

Acknowledgements. The *Apple University Consortium (AUC)* provided five *Apple iPad 2* kits for the phase one study as part of their *Seeding Equipment* program for member universities. This research is funded by a *Queensland Government Smart Futures Fellowship*, and co-sponsored by *National ICT Australia (NICTA)*. NICTA is funded by the *Australian Government* as represented by the *Department of Broadband, Communications and the Digital Economy* and the *Australian Research Council* through the *ICT Centre of Excellence* program.

References

1. Ek, K., Söderholm, P.: The devil is in the details: Household electricity saving behavior and the role of information. Energy Policy 38(3), 1578–1587 (2010)
2. Foster, D., Lawson, S., Blythe, M., Cairns, P.: Wattsup?: motivating reductions in domestic energy consumption using social networks. In: Proceedings of NordiCHI, pp. 178–187. ACM (2010)
3. He, H., Greenberg, S., Huang, E.: One size does not fit all: applying the transtheoretical model to energy feedback technology design. In: Proceedings of CHI, pp. 927–936. ACM (2010)
4. Mankoff, J., Fussell, S., Dillahunt, T., Glaves, R., Grevet, C., Johnson, M., Matthews, D., Matthews, H., McGuire, R., Thompson, R., et al.: Stepgreen.org: Increasing energy saving behaviors via social networks. In: Proceedings of ICWSM, vol. 113 (2010)
5. Marcus, A.: Dashboards in your future. Interactions 13(1), 48–60 (2006)
6. Nisi, V., Nicoletti, D., Nisi, R., Nunes, N.: Beyond eco-feedback: using art and emotional attachment to express energy consumption. In: Proceedings of C&C, pp. 381–382. ACM (2011)
7. Paulos, E., Kim, S., Kuznetsov, S.: The rise of the expert amateur: Citizen science and microvolunteerism. In: From Social Butterfly to Engaged Citizen: Urban Informatics, Social Media, Ubiquitous Computing, and Mobile Technology to Support Citizen Engagement, p. 167 (2011)
8. Rowley, J.: The wisdom hierarchy: representations of the DIKW hierarchy. Journal of Information Science 33(2), 163 (2007)
9. Shneiderman, B., Plaisant, C.: Designing the User Interface: Strategies for Effective Human-Computer Interaction, 4th edn. Addison Wesley (2004)
10. Spagnolli, A., Corradi, N., Gamberini, L., Hoggan, E., Jacucci, G., Katzeff, C., Broms, L., Jönsson, L.: Eco-feedback on the go: Motivating energy awareness. Computer 44(5), 38–45 (2011)
11. Strengers, Y.: Designing eco-feedback systems for everyday life. In: Proceedings of CHI, pp. 2135–2144. ACM (2011)
12. Vande Moere, A., 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)
13. Yates, S., Aronson, E.: A social psychological perspective on energy conservation in residential buildings. American Psychologist 38(4), 435 (1983)

An Activist Lens for Sustainability: From Changing Individuals to Changing the Environment

Eva Ganglbauer, Wolfgang Reitberger, and Geraldine Fitzpatrick

Vienna University of Technology, Human Computer Interaction
Argentinierstrasse 8, Vienna, 1040, Austria
{eva.ganglbauer,wolfgang.reitberger,geraldine.fitzpatrick}
@tuwien.ac.at

Abstract. Design for sustainability is of much interest in Persuasive Technology (PT) and interventions are often targeted to change individual's behaviour. These interventions aim to change lifestyles to be more ecologically sustainable, however the social and economic circumstances individuals live in often counteract these intentions. Activism has been proposed as a way to address such social circumstances. The contribution of this paper is to further develop an activist lens to present strategies for interventions that address policy makers as well as provide insight into how individuals can engage in activism supported by technology to advance change. Our activist lens points to active data generation and perceived agency by individuals and hybrid forms of interventions. We also address the limitations of technology in such approaches. An activist lens on sustainability and PT might provide a useful new entry point for designing change interventions from the individual to the collective.

1 Introduction

In recent years sustainability has become a prominent research topic in Persuasive Technology (PT) and Human Computer Interaction (HCI). Early research in PT was mostly concerned with the use of interactive systems to influence individuals towards more sustainable behaviour [9]. These systems commonly used technology to infer and present feedback about an environmentally relevant behaviour such as sustainable mobility and energy consumption and aimed to make the user aware of the environmental impact of their actions. The raised awareness should ideally lead to more pro-environmental behaviour as a result. However, the sole focus on changing individual behaviour is controversial [2] and is only one approach in the much larger space of possibilities for change towards more sustainability. An equally important aspect is influencing the circumstances, and the decision makers who control them. For example, Mankoff et al. [8] point out that the individual level is only one area to target for sustainability through design, and is complemented by efforts directed at the societal level. Woodruff et al. [12] also conclude that design efforts could be directed at changing circumstances which often hinder or prohibit sustainable practices on an individual level. These positions provide us with a starting point to more deeply engage with an activism lens. This shift is important because going beyond individual

S. Berkovsky and J. Freyne (Eds.): PERSUASIVE 2013, LNCS 7822, pp. 63–68, 2013.

behaviour change would help to shift perspectives from blaming individual shortcomings without concern for context, towards realizing and acknowledging the role of the larger context and the societal circumstances play in shaping the individual's practice.

These circumstances can promote, hinder or completely disable sustainable practices. Whereas individual persuasion emphasizes what Latour terms the *intrasomatic* aspects such as raising awareness, in this paper we focus on an activist approach aimed at influencing the *extrasomatic* reality or circumstances [7]. As our consciousness is determined by our society and social being, it is the extrasomatic domain which in turn will have effects on the collective consciousness and affect practices on an individual level. Even the best designed and most well intended PT application to foster sustainable behaviour cannot persuade users to engage in the desired behaviours if the circumstances are not allowing or supporting them. At this point it is better to advocate change of the extrasomatic social and political circumstances and support individuals to engage with activism.

2 An Activist Lens

Activism is defined as taking actions to promote social, political, economical and environmental change. Generally people need to be discontented about certain circumstances to become agents of change. In this paper we use a definition of an activist as not only a person who engages in short term physical protest, but one who takes any form of action, ranging from digital to physical acts with the aim to not only to change individual consumer behaviour, but also to change circumstances at whatever level of scale is most relevant. PT can support activists with information and communication technologies (ICTs) on an individual level, but also support communication and cooperation among individuals for collective action.

Within HCI some work has already addressed pro-environmental citizen sensing as one particular strand of the much larger repertoire of activism. Aoki et al. [1] explored data from air quality sensors on public street sweepers to reflect on the role this data can have in pro-environmental discourses with stakeholders. Kuznetsov et al. [5] researched participatory citizen sensing and how participants engage with space, the placement of sensors and the use of the collected data for broadcasting, sharing and activism. In this paper, instead of focusing on specific strands of activism like citizen sensing, we provide an analysis of activism and ICTs, resulting in a repertoire of activism, and in turn identify strategies for interventions. Thus we engage with literature from fields of sociology, communication and political science that address these issues. In the following section, we provide an overview and critical understanding of activism and in turn reflect on the potential strategies for PT in this domain.

2.1 A Repertoire of Activism

There already exists some work that describes different dimensions for categorization that can be useful for our purposes. Van Laer & Van Elst [11] describe different '*action repertoires*' and use the dimensions of 'high vs. low thresholds' for

participation and 'ICT-supported vs. ICT-based' actions. *Consumer behaviour, donating money* and *demonstrations* are described as repertoires with a low threshold and supported by ICTs. *Occupations* or *transnational demonstrations* are also supported by ICTs, but have a high threshold for participation. There are as well repertoires solely based within ICTs, such as *online petitions* with a low participation threshold and more based on the individual, whereas *culture jamming*[1]*, email-bombs*, *protest websites* or *hacktivism*[2] have higher thresholds to participate [11]. We also identify three other instances of action repertoires not explicitly mentioned by Van Elst and Van Laer, which can also be important for HCI and PT, namely citizen sensing, grassroots movements and open data initiatives.

With new advances in low-cost sensor technology *citizen sensing* is a design area where HCI can engage, as seen in the previous discussion. Citizen sensing is the act of collectively sensing environmental data to share with others. Such collections can be seen as persuasive 'social currencies' or techno-political tools, data that can be used to share and negotiate collectively with policy makers [1, 5]. The bottom-up approach of citizen sensing can change power relations in empowering citizens to act as agents of change. Volunteer-driven, non-hierarchical organizational models characterize bottom-up forms of activism such as *grassroots movements.* These are often lacking funding, spontaneous by nature and demand ICTs and other methods of support [6]. Generally grassroots movements are supported by ICTs and have a low threshold in participation if online [10] and a higher threshold when the movement crosses over into the offline world. Change does not necessarily always come from the bottom. Data can also be provided by governments through *open data initiatives* to raise new potential for participation and activism starting with a top-down approach. The democratization of data gave rise to many initiatives [4] that provide data to the public for free. These initiatives can potentially enable citizens to participate in decision-making because they can process and present data in creative forms, and allow decision makers to understand this data to anticipate next steps.

2.2 ICTs for Increased Activism?

ICTs are not necessarily a panacea for social action and here we take a more critically reflective stance on the role of ICTs for a repertoire of activisms. In particular we identify two general pitfalls that lead to a more nuanced understanding of the potential of activist interventions. The first issue is about if, and to what extent, technology promotes social change: "*Evidence that ICT use is producing significant social change does not mean that the changes identified are inherent to the technology. Used in different contexts, technologies yield different effects.* [3, p. 217]". Without question, forms of activism have changed along with the technological tools and actions now can be more global, and be mobilized more quickly. But the data on this is open to dispute, with some claiming that ICTs increase participation, while others

[1] A form of resistance to disrupt big mainstream institutions e.g. by subverting corporate advertising campaigns.

[2] The use or attack of ICT infrastructure in pursuit of political protest that may be legal or illegal.

finding no effects [3]. What the literature suggests though is that an ICT intervention alone cannot trigger activism, as it is mainly people who drive transformations. If people just engage in the online world, a weak spot of ICT-based activism can be that the engagement is just online, referred to as slacktivism or clicktivism, i.e., activism that makes it very easy for users to participate online and to declare one's opinion but not engage in the activist movement offline.

The second issue concerns the *digital divide*, with those on the wrong side of the divide not able to participate in activism that is supported by, or based in, ICTs. A gap between ICT users can be observed globally, with more ICT users in industrialized countries and fewer in less industrialized countries. The digital divide also applies within nations as the socially underprivileged tend to have poorer access to, and lower skills to engage with ICTs and are therefore more likely to be excluded from participation and activism.

3 To Design with an Activist Lens

We have engaged with literature from the fields of sociology, communication and political science to gain an understanding of a repertoire of activism and the potentials of ICTs around activism. For PT, this has several implications. At points here, we will use cycling as one pro-environmental practice to illustrate some of the following as intervention strategies.

From Individual to Collective: While individual PT approaches can bring about individual change, it can also be useful to enable individuals to instigate and manage collective action and grassroots movements. This would transcend the closed feedback loop of a behaviour change intervention and empower users to transform themselves from being someone to be changed by a system to be active agents of both local and systemic change. For social movements PT designers can consider both how to support mobilization, and also if opportunity structures exist where ICTs can foster collective activism. A cycling app could not only target to engage people in cycling activities but also remind them about protest bike rides such as critical mass[3] or connect people with similar interests for making changes.

Active Data Generation: In addition to citizen sensing and personal feedback technologies there is a broad range of further opportunities for activism. It is also possible to engage people to actively collect data for decision makers and other citizens, where active rather than passive collection [1, 5] can also more likely promote opportunities for individual awareness and reflection, e.g. cyclists augmenting a map by pinning dangerous spots for local government intervention and re-construction of infrastructure. Active data generation can act as a collective filter to add a rich qualitative layer of information for stakeholders and highlight particularly relevant data points by the crowd. If data is generated actively, people can decide on their own how they want to shape information for decision makers, whereas with automatic sensing the data is evaluated based on pre-defined filters.

[3] Monthly protest bike event held in 300 cities around the world that aims to communicate problems to city council and/or reclaim the streets.

Hybrid Forms: There is also potential to combine these strategies. Even traditional approaches of behaviour change and PT can be integrated with activist repertoires to create new *hybrid forms* of ICT interventions. From a design perspective the intra- and extrasomatic can feed and inform each other or be engaged with in parallel. For example, traditional behaviour change interventions with personal benefits can be integrated with a repertoire of activist approaches as noted previously. For example, cyclists could be using an app that both tracks their activity levels and captures air quality data and this data can contribute to open data platforms adding a layer of real time information. Or a green mobility app could include the aforementioned pinning capability and allow cyclists to generate data for improving the cycling infrastructure. Such behaviour change interventions can be re-conceptualized to empower users by providing an interaction with responsible stakeholders.

Perceived Agency*:* In order to generate active or passive data for decision makers, or to set up grassroots movements, the participants' belief in the *possibility of change* is crucial and ICT can play a role in promoting such visibility. Feedback about the activist intervention can be made visible to the contributors, to enable them to assess whether desired outcomes have been achieved. It is important for users not only to perceive being able to influence/control their own behaviour, but also to enable them to influence areas that they are concerned with. Thus ICTs can act as a mediator to expand influence from the personal to the societal and political sphere. Furthermore ICTs can help to discover opportunities for change within the personal sphere that can be directly controlled by a person, again increasing the perceived and real agency.

4 Conclusion

We have identified a repertoire of opportunities to engage with activism in addition to existing sustainability approaches. The aim is to trigger researchers and practitioners in PT and HCI to think beyond individual behaviour interventions as the main tool for change. We have opened the design space to embrace activism as an additional means to also address broader societal transformation. To this end we have proposed an expanded set of possibilities for change interventions using an activist lens. These promote consideration of: the individual to the collective; active data generation; hybrid forms of interventions; and perceived agency. Our intention is not to advocate activism through ICTs as a panacea but to add a more nuanced perspective for the PT and HCI community by taking an activism lens. While it is crucial to make it as easy as possible for people to engage in activism through ICTs, there exists the potential for slacktivism or clicktivism as previously noted. On the other hand the researchers and designers can take the opportunity and design interventions that can support mechanisms not only to advocate individual behaviour change, but to also change local and social circumstances or both through hybrid forms, depending on what is within the personal or collective sphere of influence or, even better, control. We hope that this contribution inspires PT and HCI researchers to embrace the activist lens when addressing change towards sustainability.

References

1. Aoki, P.M., Honicky, R.J., Mainwaring, A., Myers, C., Paulos, E., Subramanian, S., Woodruff, A.: A vehicle for research: using street sweepers to explore the landscape of environmental community action. In: Proc. CHI 2009, pp. 375–384. ACM (2009)
2. Brynjarsdottir, H., Håkansson, M., Pierce, J., Baumer, E., DiSalvo, C., Sengers, P.: Sustainably unpersuaded: how persuasion narrows our vision of sustainability. In: Proc. of CHI 2012, pp. 947–956. ACM (2012)
3. Garrett, K.R.: Protest in an Information Society: a review of literature on social movements and new ICTs. Information, Communication & Society 9(2), 202–224 (2006)
4. Kalampokis, E., Hausenblas, M., Tarabanis, K.: Combining Social and Government Open Data for Participatory Decision-Making. In: Tambouris, E., Macintosh, A., de Bruijn, H. (eds.) ePart 2011. LNCS, vol. 6847, pp. 36–47. Springer, Heidelberg (2011)
5. Kuznetsov, S., Davis, G., Cheung, J., Paulos, E.: Ceci n'est pas une pipe bombe: authoring urban landscapes with air quality sensors. In: Proc. CHI 2011, pp. 2375–2384. ACM (2011)
6. Kuznetsov, S., Odom, W., Moulder, V., DiSalvo, C., Hirsch, T., Wakkary, R., Paulos, E.: HCI, politics and the city: engaging with urban grassroots movements for reflection and action. In: Ext. Abstracts CHI EA 2011, pp. 2409–2412. ACM (2011)
7. Latour, B.: Where Are the Missing Masses? The Sociology of a Few Mundane Artifacts, pp. 225–258. MIT Press (1992)
8. Mankoff, J.C., Blevis, E., Borning, A., Friedman, B., Fussell, S.R., Hasbrouck, J., Woodruff, A., Sengers, P.: Environmental sustainability and interaction. In: Ext. Abstracts CHI 2007, pp. 2121–2124. ACM (2007)
9. 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)
10. Schumann, S., Klein, O., Douglas, K.: Talk to Act: How Internet Use Empowers Users to Participate in Collective Actions Offline. In: Bang, M., Ragnemalm, E.L. (eds.) PERSUASIVE 2012. LNCS, vol. 7284, pp. 79–89. Springer, Heidelberg (2012)
11. Van Laer, J., Van Aelst, P.: Cyber-Protest and Civil Society: the Internet and Action Repertoires in Social Movements. In: Handbook of Internet Crime. Willan Pub., Cullompton (2010)
12. Woodruff, A., Hasbrouck, J., Augustin, S.: A bright green perspective on sustainable choices. In: Proc. CHI 2008, pp. 313–322. ACM (2008)

Improving the Design of Online Applications for Social Benefit through a Behaviour Change Model

Kathryn Gough and Jillian Hamilton

Faculty of Creative Industries, Queensland University of Technology, Brisbane, Australia
{kathryn.gough,jg.hamilton}@qut.edu.au

Abstract. Community support agencies routinely employ a web presence to provide information on their services. While this online information provision helps to increase an agency's reach, this paper argues that it can be further extended by mapping relationships between services and by facilitating two-way communication and collaboration with local communities. We argue that emergent technologies, such as locative media and networking tools, can assist in harnessing this social capital. However, new applications must be designed in ways that both persuade and support community members to contribute information and support others in need. An analysis of the online presence of community service agencies and social benefit applications is presented against Fogg's Behaviour Model. From this evaluation, design principles are proposed for developing new locative, collaborative online applications for social benefit.

1 Introduction: The Design Problem

Community support agencies routinely employ a Web presence to provide information to people in need. This undoubtedly increases access to local services, events, advice and support. However, because such agencies tend to offer diverse, geographically dispersed services, and provide information as separate entities, it can be difficult for those facing hardship to locate individually appropriate support. This issue was identified during research into ways to support and extend the digital strategy of a local agency, *The Exchange,* in Brisbane Australia. As one community worker put it,

> *One of the challenges people have when experiencing need is finding appropriate services. [Our] agency frequently gets phone calls from people who are looking for a specific service and have often made many calls before reaching this agency.* [Interview, 16 May 2012, Brisbane]

Information sharing is crucial to helping people in need. As Price-Robertson and Knight (2012, 6) argue, "[When] organisations, services and community members … work interactively and transparently to ensure correct information is communicated in a timely and effective manner [they become an] enabler of community resilience". Emergent technologies such as locative media can help by facilitating the mapping of geographical (as well as service provision) relationships between agencies, while aggregation tools can assist by combining information sources, and content management systems (CMS) can enable information updating. However, while the mapping and

S. Berkovsky and J. Freyne (Eds.): PERSUASIVE 2013, LNCS 7822, pp. 69–74, 2013.

aggregation of information is undoubtedly beneficial, it does not fully exploit the potential of these new technologies unless it also activates the latent goodwill of the community. As community workers who participated in the recent Brisbane research suggested, while mapping information on services and events is important, facilitating two-way communication, information sharing, and collaboration with the community can enable support agencies to reach further and be more effective [Interview, 16 May]. Community members are a potentially rich source of information and support and tapping into the information and resources they have to offer through strategies like crowd-sourcing can help build social capital.

This paper presents a case study in the form of preparatory research for a locative media tool for social benefit. Called Upraxia (a Greek term denoting the realization of well-being), the goal of the application is to enable community support agencies, their clients and community members to collaborate to produce a shared, locative directory of information and services, community-generated advice, and support for people in need. Establishing design principles for this application has involved a participatory approach: working with a community agency to establish their goals and how new technologies can support them as well as employing the analytical framework of persuasive media to evaluate current online strategies of local service agencies, and their capacity to achieve these goals.

2 Analysis of Stakeholder Needs

In order to ascertain the goals of community support agencies and how they might be enabled by emergent technologies, a local, representative support agency called *The Exchange* was consulted. Open-ended interviews focused on their current online and offline strategies, their goals, and opportunities for deploying technologies to achieve them. It established that the primary goal of the agency is to provide services to socially disadvantaged local people. Providing better access to information on services and events is a key priority. They see the benefit of providing locative information, commenting that, "*because clients can be restricted to access[ing] services where they live, mapping capabilities would be beneficial.*" This mapping, they suggested, should include their own services as well as those of others in their local area because, as yet, there is no centralised information resource.

Another key priority of the agency is to be proactive in information sharing. They stated that, "*If certain information stays secret with certain people within certain organisations then its not empowering people."* Moreover, they are keen to support the community to contribute to this information. However, they emphasized the importance of making sure the application provides timely, accurate and verified information, so primary advice and links to resources should be contributed by the agency. Another concern involved privacy risks, as their clients are often unaware of online security. While the initial interviews focused on these general needs of the agency, a follow-up questionnaire with checkboxes and short answers focused on the types of services to be mapped and potential opportunities for community members to be involved in communication provision. They prioritised supporting the community to post information, comment, send private messages, save posts, and subscribe.

This needs analysis helped to establish an initial set of goals for a new application. They include expanding information sources beyond the agency by aggregating and locatively mapping services available in the local area and empowering the community to collaborate in information and support provision by facilitating a two-way conversation, while providing safeguards to ensure accurate information and privacy.

3 Evaluating the Online Presence and Tools of Support Services

An analysis of existing community support agency sites was conducted against the objectives of aggregating information, mapping services and information sharing. Particular consideration was given to the goal of encouraging and motivating community participation. The field of persuasive media provides a useful analytical framework for determining the capacity of technology systems to persuade people to act (in this case to contribute). Designers influence and persuade in many ways through conscious decisions about the way information is presented in terms of colour, space, and layout. However, as Fogg (2002) argues, technologies are potentially potent tools of Mass Interpersonal Persuasion, which can bring about positive changes in domains as diverse as health, business, safety, and education if they are envisaged as more than a medium for information delivery. Interactive systems can function as social actors that provide experiences, which shape behaviour through the way in which user interaction is framed, sequenced and structured. Fogg's formulation of captology provides a "systematic way to think about the factors underlying behaviour change"–from inducing new one-of behaviours such as signing up to an online mailing list, to increasing a specific behaviour such as exercising more each day (2009, 1). An effective, technologically-mediated experience is contingent upon making users feel comfortable about making decisions and helping them to act, while motivating behaviour change is contingent upon providing appropriate triggers, motivation and ability to act.

Let's now consider the online presence of support agencies against the goals of the Upraxia system and Fogg's Behaviour Change model. Red Cross Australia, Mission Australia and Communify Queensland all focus on informing site visitors about the organisation's services, campaigns, news and event, through concise descriptions on a homepage. All ensure ease of access to information through clear layout and fundamental usability principles (as identified by Sharp, Rogers, and Preece, 1994). All three sites have embraced social media, with links and icons directing visitors to Facebook pages and Twitter feeds of the agencies. In the case of Communify Queensland, social media icons appear below every event and activity to encourage sharing. And, because each encourages community engagement in activities and donations, all of the sites provide a call to action. Red Cross Australia and Mission Australia encourage donations through a prominently located, bright clickable button on the landing page and subsidiary pages. Attention has been given to ensuring the ability to perform this action efficiently. The donation process of the Red Cross site, for example, involves three simple, automated steps, while Mission Australia and Communify Queensland also have an efficient one-page signup process for an 'eNews' mailing list.

If we consider these sites in relation to Fogg's captology and his recommendation to provide triggers, increase the motivation to act, and facilitate the ability to contribute,

then we might say that all of these objectives are realised at some level. Motivation to act is provided through information provision (in the expectation that greater understanding will lead to trust, empathy, commitment and motivation). And attention to usability and a smooth, streamlined experience, these sites ensure the ability to act.

However, information provision is the only tool these websites employ to trigger interest and involvement. While social media is incorporated, it is primarily designed to increase the dispersal of information, and the call to action of two of the sites–to sign up for a newsletter–simply increases information saturation. An 'interactive' experience is provided, but it is simply navigating between information, or committing to donation. As Fogg (2010, 11) argues, new (what he calls Green) behaviours can potentially be induced by simply increasing awareness of the desired behaviour and providing triggers. This means that seemingly straightforward design decisions, such as the strategic placement of a button on a website, can provide a path to a desired action. However, while the principles of usability, effective content presentation, and enabling action through easy to follow steps are no doubt valuable, in the end this is a model of passive information reception. The capacity that the community service stakeholders (users and the agencies) interviewed in this study called for–such the aggregation of services, relational mapping to other agencies, effective service provision (alongside information provision), and the social aspects of contributing and community collaboration are not evident. The latent potential of web, locative and networking technologies is in providing the capacity for agencies to consolidate their collective presence, to stimulate community engagement, and to build social capital.

4 The Potential for New Online Strategies for Social Benefit

There are examples of social benefit applications that have taken steps in the directions identified by the user needs analysis. MyCommunityDirectory, for example, is an Australian online application that collates community service information. Agencies can use it to post and update listings to promote their services. After creating an account and establishing a profile, a maintenance phase involves updating on a regular basis. The application combines the contributions to display a range of information on support options. In this way, it facilitates distributed content provision, the co-presentation of services, and relational representation of a network of agencies. From the perspective of a person in need of support, this approach is no doubt invaluable as it provides organised information on a range of options. However, MyCommunityDirectory does not include locative functionality. There has been a rapid proliferation of locative media applications, which deploy 3G mobile phones and geo-mapping technologies. However, as de Souza e Silva and Sutko (2011, 38) argue, "[locative media requires designers to] rethink the relationship among users, spaces, information and interfaces." If this is achieved, locative media can visualise the geographical relationships between services, and it can also assume a social dimension with connections within local communities. At present, MyCommunityDirectory provides dislocated information distribution.

There are social benefit applications that enlist community involvement and co-operation in local problem solving. Yoink is a collaborative application that uses a

Facebook API to encourage people to post information about available and needed goods. Its commitment phase involves linking the app to a Facebook account, while maintenance involves periodically posting offered or wanted goods. This serves to trigger responses by others. Ability is ensured through the use of the Facebook interface, which is already familiar to most users. And motivation is provided through the users social graph. In a paper on Mass Interpersonal Persuasion, Weiksner and Liu and Fogg (2008) argue that including a structure or process through an expanded social graph, and the (measured) display of impact, serves to encourage a desired behaviour through the endorsement by trusted peers. This application no doubt provides benefits to people in need by harnessing social capital and community goodwill. However, it does have limitations. While the intention is for localised exchange, this application does not use locative functionality. It is restricted to 'wall' functionality, private messaging and a forum. In addition, linking requests with Facebook information may deter participation, due to privacy concerns. Trust is an issue that must be considered in the provision of online services and information sharing and users can be assured that information is accurate, timely and respects privacy through moderation and verification, which is not provided here.

Givit, an Australian online portal, also facilitates the donation of goods to people facing hardship. It solves potential privacy concerns by acting as an intermediary, posting requests on behalf of charities. The Givit community then responds, providing what is needed to the charity to distribute. Givit implements what Fogg refers to as Green Path behaviours, which lead from an initial investigation phase, in which the site visitor seeks information and assurance; to a transitional phase which involves securing commitment and typically requires a one-off, low cost action which takes into account limited ability to perform novel tasks; to long-term commitment, where the behaviour becomes routine (Fogg 2010, 3). In this final phase, cyclical, predictable prompts (or irregular cues) provide triggers to act (Fogg 2010, 23, 24). This approach requires different forms of triggers in each phase. In the case of Givit, a large email form field that appears on the homepage in the transitional phase, ensuring the ability to perform a low cost, behaviour (signing up for a newsletter). The maintenance phase involves the receipt of a weekly list of requests, which acts as a 'hot trigger' by highlighting need 'in the moment'. This serves to make the recipient and their hardship 'real'. Fogg argues that designers can implement behaviour change by putting hot triggers in front of motivated people (Fogg 2010, 25). Givit achieves this through a regular cycle of triggers, which emphasize currency and urgency to those committed to help.

5 Design Principles for a Social Benefit Application

In *Design Activism* (2009, 187), Fuad-Luke argues that, for 200 years, "design was self-absorbed [and] besotted in the power bestowed upon it by commercial interests and assured of its ubiquitous presence in consumers' lives." Designing online resources and tools for social benefit provides the opportunity for some redress. However, to be effective, we must adopt some key design principles. We have set out to establish these here because new technologies such as locative media do not bring

with them the precedents and principles of other, longer established design fields. By seeking the input of a local service agency, undertaking a contextual review of community support agency web sites, and analysing early social benefit tools against the framework of persuasive media, the principles we have established for Upraxia as well as other similar social benefit applications can be summarised as follows.

Firstly, applications for social benefit must prioritise the importance of information provision as an "enabler of community resilience" and, to this end, information must be timely, appropriate and reliable. Secondly they can optimise users' access to services by aggregating and contextually locating and relating them. New technologies, such as locative media, can help to do this by producing geographic visualisations and comparisons that are local and focal. Thirdly, to be effective, it is imperative that new applications are designed in ways that empower support services and the community to co-operatively support people in need. To do so, we must move beyond information delivery and design interactive experiences that optimise the capacity for local communities to contribute and build social capital. That is, we must ensure the ability to act, while ensuring risk minimisation. We must also establish the motivation to act, by strategies such as conveying immediate need, normalising the desired behaviour through a social graph or the visualisation of positive outcomes. Networking tools can be harnessed for this task, but when contributing behaviour is new, designing phased commitment levels with a stepped sequence of triggers allows us to go beyond simply increasing awareness of a social issue, to providing an experience that prompts a contribution to supporting those in need.

References

1. Bruns, A.: Social Media vs the Floods, http://cci.edu.au/about/media/social-media-vs-the-floods
2. Fogg, B.J.: Persuasive Technology: Using Computers to Change What we Think and Do. Morgan Kaufmann Publishers, Amsterdam (2002)
3. Fogg, B.J.: A Behavior Model for Persuasive Design (2009), http://www.bjfogg.com
4. Fogg, B.J.: Green Path Behaviour Guide. Stanford Persuasive Technology Lab. (2010)
5. Fuad-Luke, A.: Design Activism: Beautiful Strangeness for a Sustainable World. Earthscan, London (2009)
6. Price-Robertson, R., Knight, K.: Natural Disasters and Community Resilience: a Framework for Support. Child Family Community Australia paper No.3 (2012)
7. Sharp, H., Rogers, Y., Preece, J.: Interaction Design: Beyond Human-Computer Interaction. John Wiley and Sons, London (1994)
8. Silva, A.d.S.e., Sutko, D.M.: Theorizing Locative Technologies through Philosophies of the Virtual. Communication Theory 21, 23–42 (2011)
9. Weiksner, G.M., 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)

Curbing Resource Consumption Using Team-Based Feedback

- Paper Printing in a Longitudinal Case Study -

Souleiman Hasan[1], Richard Medland[2], Marcus Foth[2], and Edward Curry[1]

[1] Digital Enterprise Research Institute, National University of Ireland, Galway, Ireland
[2] Urban Informatics Research Lab., Queensland University of Technology & NICTA, Australia
{souleiman.hasan,ed.curry}@deri.org,
{richard.medland,m.foth}@qut.edu.au

Abstract. This paper details a team-based feedback approach for reducing resource consumption. The approach uses paper printing within office environments as a case study. It communicates the print usage of each participant's team rather than the participant's individual print usage. Feedback is provided weekly via emails and contains normative information, along with eco-metrics and team-based comparative statistics. The approach was empirically evaluated to study the effectiveness of the feedback method. The experiment comprised of 16 people belonging to 4 teams with data on their print usage gathered over 58 weeks, using the first 30-35 weeks as a baseline. The study showed a significant reduction in individual printing with an average of 28%. The experiment confirms the underlying hypothesis that participants are persuaded to reduce their print usage in order to improve the overall printing behaviour of their teams. The research provides clear pathways for future research to qualitatively investigate our findings.

1 Introduction

Consumption of resources is a major concern from financial and environmental perspectives. Paper printing, electric heating and transportation are all examples of resource consumption related to everyday activities in normal business environments. Modern computing technologies and devices to track such resource consumption can intervene to curb wastage. However, proper intervention methods are yet to provide a significant impact [1]. There is a need to support such methods with theoretical and empirical studies involving users and consumers of resources.

Medland [2] devised a feedback-based intervention method to reduce wastage in paper printing. The method is based on providing staff in an office environment with their printing statistics. Within the approach, multiple feedback "flavours" were evaluated, ranging from traditional usage statistics, eco-metric printing statistics to comparative and competitive statistics. In this paper we build on this approach by using a team-based feedback method that leverages the membership of individuals in local teams with no individual performance being communicated.

S. Berkovsky and J. Freyne (Eds.): PERSUASIVE 2013, LNCS 7822, pp. 75–86, 2013.

The hypothesis is that by providing people with their team's resource consumption performance in relation to that of other local teams, a social incentive is created for individuals to improve their team's overall performance. We tested the hypothesis in an office environment for over a year, with paper printing as the particular instantiation of resource consumption. Our experiment includes 16 participants belonging to four teams. It compares the printing performance of participants during a baseline period before applying – or communicating intent to apply – the feedback method, with their performance after applying the method. Results show a statistically significant improvement in individual printing behaviour after applying the team-based feedback method confirming the tested hypothesis.

The rest of this paper is organised as follows. Section 2 reviews related work in the areas of persuasive technology, social norms and feedback-based pro-environmental behaviour. Section 3 describes the case study environment and the software architecture employed to implement the experiment and then details the research design. Section 4 presents the results, and Section 5 discusses these results in light of three themes. Section 6 concludes the paper with suggested prospects for future research.

2 Related Work

2.1 Persuasive Technology

Our research employs persuasive technologies that increase awareness of resource consumption providing pathways for restructuring behaviour. According to persuasive technology theorists such as Fogg [3, 4], "… a persuasive technology is fundamentally about learning to automate behaviour change" (p. 1). These technologies are designed to persuade those using them through a range of methods, particularly social influence, to change their behaviours by first changing their attitudes. Lockton et. al. [5] extend this definition by including situations, such as safety systems, where attitude change is not a precursor for behaviour change, and is in fact not a consideration. Beyond these considerations - and given our focus on teams in this research - Khaled et. al. [6] point to a need for consideration of collectivist versus individualist cultural dimensions. As technologies become more pervasive they also often need to persuade, if they are to assist us in our daily lives [7]. Speaking to this conclusion were researchers such as Davis [8] who considered the perceived usefulness and perceived ease of use of ICT as determinants for user acceptance of those technologies. What has become more obvious in the last decade is that as the uses and users of ICT expand, so do the expectations of the user experiences when using that technology [9].

2.2 Social Norms

Our research takes advantage of social norms and the well-established roles they play in predicting behaviour [10, 11]. For the purposes of our research, social norms were divided into two types: descriptive and injunctive. *Descriptive* norms provide information on commonly accepted behaviours; the standard that people attempt to adhere

to [10]. Deviant behaviour diverges from this norm and is often negatively perceived. *Injunctive* norms are shaped by cultures, providing guidance on behaviours that are approved or disapproved of. Applying descriptive or injunctive normative information in isolation can cause reflexive results, such as causing outliers to converge, even when they are exemplars of the behaviour you wish to encourage. To counteract this "boomerang" effect, the descriptive and injunctive feedback should be applied in tandem, providing reinforcement for high-achievers, whilst communicating the need for change to underperformers [11, 12].

2.3 Feedback and Pro-environmental Behaviour

Feedback is inherently information and is often available, and linked to, contextually relevant events [2]. We know that information provided at these moments - or at regular intervals - has attributes or qualities that aid in understanding, accessibility, and retention (such as being memorable or surprising) [13]. For most of the resources consumed as part of daily living, feedback in some form is available, though often fairly simplistically [14]. This simplicity leads to a common issue: translating the consumption data into readily accessible - and actionable - information for the individual [13]. More problematic still, even though some individuals understand their resource consumption, they do not take up relevant pro-environmental behaviours to reduce it [15]. In the setting for our research paper was a monitored and digitally recorded resource, despite this, the best estimates for the individual prior to our study was the visual assessment of the paper to be recycled. Our current research leverages these findings to provide regular accurate feedback in a persuasive manner for paper using a number of metrics.

3 User Study

In order to study the effect of team-based feedback on individual consumption behaviour, we chose to set up an experiment to track printing practices in an office-based environment that natively has different working units operating as information-intensive semi-autonomous teams.

3.1 The Case Study Environment

The experiment was conducted within the Digital Enterprise Research Institute (DERI) located in Galway, Ireland[1]. DERI is a research institute focusing on research activities in multiple aspects and applications of the Semantic Web [16]. As of November 2012, DERI employed about 130 staff. Staff can be grouped into: research interns, master and PhD students, post-doc researchers, research assistants, research fellows, senior research fellows, professors, technical staff, and administrative staff. Staff is divided among a set of about 20 organisational units. Each of the units

[1] http://www.deri.ie/

conducts research in an aspect of the Semantic Web. A set of communal printers are distributed around the DERI building. Reasons for printing in DERI are summarised in the following categories:

- Administrative printing by administrative and research staff.
- Printing research proposals, theses, academic and technical reports.
- Printing academic papers for internal review and reading purposes.

3.2 Print Monitoring

Printing is tracked in the DERI building using print monitoring software. The monitoring software installed on the main print server in DERI monitors the print queues of all the existing printers connected to the DERI network. It records each print job with various attributes such as date and time, and the number of printed sheets. This data gathering process is invisible to staff and some incidents occurred during our study demonstrating that participants were unaware of the data recording.

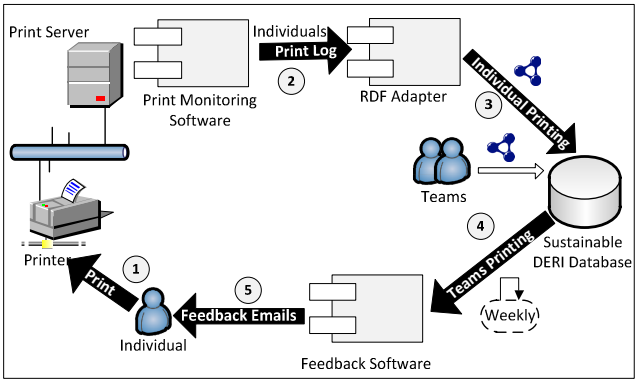


Fig. 1. Feedback cycle architecture

The software can export detailed print logs for any date and time range. However, these logs do not associate staff with their business units or teams. As a proxy, we employed a graph database that uses the Semantic Web Resource Description Framework (RDF) [17]. It consolidates organisational and various sustainability-related entities in DERI in one place based on Linked Data principles [18]. It is regularly fed by various data collectors including an RDF adapter to print-monitoring software developed under the Sustainable DERI project [19]. The Sustainable DERI database can be queried using the Semantic Web SPARQL query language [20]. The Sustainable DERI Linked Dataspace is investigated in more details in [21] with a feedback scenario for cloud computing consumers in [22].

3.3 Feedback Software

The feedback software used for the experiment builds on previous work [2]. The feedback software is responsible for querying the Sustainable DERI database on a weekly basis in order to generate and send emails to individual participants. Email was chosen as the communication channel as it is deemed legitimate by all staff (where others such as instant messaging are not). Figure 1 illustrates the architecture used to generate and send the feedback emails.

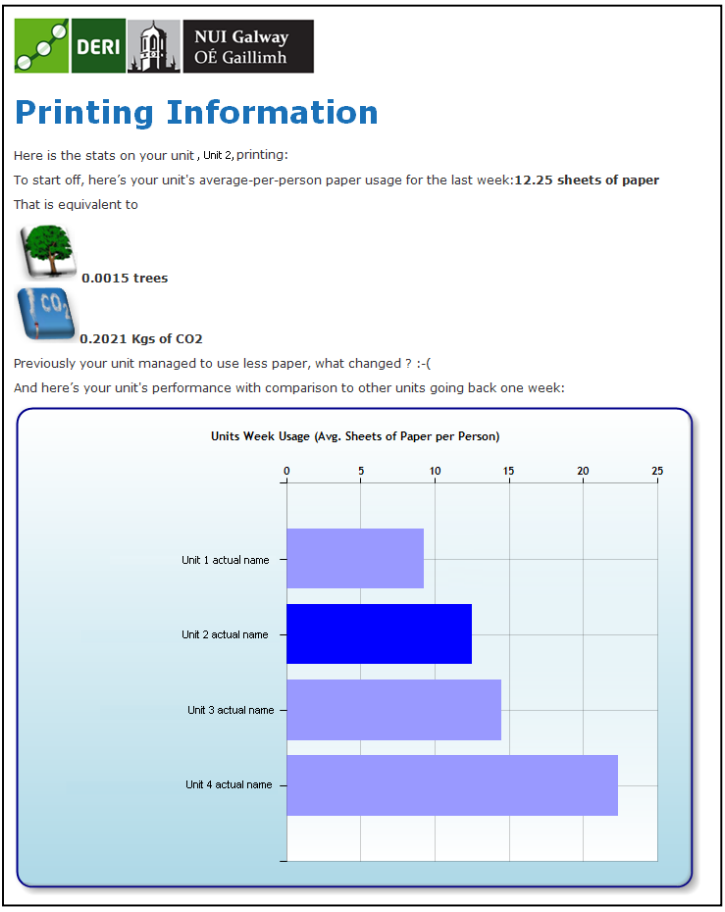


Fig. 2. An example feedback email

The email sent includes three feedback methods to communicate the same information as a previous study [2] showed that multiple methods are needed to address the multiple feedback flavours that different staff would find persuasive for them. The feedback methods are described below. To facilitate this discussion, please refer to Figure 2 for a sample email.

- **Traditional Statistics:** where the email states the number of pages printed by the person's team averaged by person. E.g. "Here is your unit's average-per-person paper usage for the last week: 12.25 sheets of paper."
- **Eco-metric Statistics:** where the number of sheets is converted to an estimated amount of trees or CO_2 emissions. E.g. "Your unit's average-per-person printing for the last week is equivalent to 0.0015 trees and 0.2021 kg of CO_2."
- **Comparative Statistics:** where a comparison is drawn on a week-to-week basis and a team-to-team basis. The comparative statistics target the use of descriptive and injunctive social norms [10, 11] in order to affect individual's behaviour. E.g. "Previously your unit managed to use less paper, what changed? ☹." In this case injunctive normative information is the sad face, denoting that in your previous week you achieved lower paper consumption. Conversely, if your efforts are better than previously recorded, a smiling face is displayed. Another example is a chart that illustrates printing performance of the participant's team in comparison with other participating teams during the previous week.

3.4 Experimental Setting

While more granular print usage is available for each participant, data was aggregated and displayed weekly. This time period forms the smallest temporal unit we deal with throughout our study. Gathering of printing data logs for all DERI members started in October 2010. Throughout the period until May 2011, no feedback emails were sent to any participants. Thus, this period forms the baseline for our study. After the end of the baseline period, team leaders were asked to participate and briefed informally. Thereafter, participants received feedback emails on a weekly basis. Participation in the experiment was voluntary. Ultimately a set of 16 members belonging to four different research units participated in the experiment as shown in Table 1. Participants started to receive feedback emails between weeks 31 and 35 of the study. Feedback emails continued to be sent until week 58 when the experiment ended.

Table 1. Participating members per teams (data anonymised)

Team	Members
Unit 1	Members 1, 2, 3,4
Unit 2	Members 5, 6, 7, 8, 9, 10
Unit 3	Members 11, 12, 13
Unit 4	Members 14, 15, 16

The main measure studied in our experiment is the total number of printed sheets of paper per person per week. The measure is evaluated using a before/after paired statistical test with baseline period as the "before" period, and the feedback period as the "after" period. However, variations might occur throughout the study in the overall printing behaviour in DERI due to project deliveries, proposal deadlines or a financial year end for example. Events such as these might cause an increase or decrease in printing by participants that is not resultant from the feedback itself, but

related to the impact of an external event. To explain this, let us assume that a person prints 10 sheets per week before applying the feedback method, and 7 sheets per week afterwards. It is not possible to safely state that there is a decrease in printing of 30% that is resultant from the feedback method. If DERI printing in general decreases from 1000 sheets per week before to 700 afterwards, then it is probable that an external variable has an effect on the amount printed. In order to isolate such external variables it was decided to use the person proportional printing over DERI printing every week. If a person prints 10 sheets in a week and DERI prints 1000 sheets during the same week, then the person proportional printing is said to be 1%.

4 Results

Our initial hypothesis was that participants would reduce their print usage in order to improve their team performance. This hypothesis would be confirmed if the print usage of each participant decreased significantly after applying the feedback method.

4.1 Individuals Level Printing Performance

Table 2 presents the weekly average printing proportions of participants for DERI as a whole. Printing data is sourced from two periods: the period before the feedback method was applied, and the period afterwards, until the end of our study.

Table 2. Participant average weekly printing proportions for DER as a whole

	Members							
	1	2	3	4	5	6	7	8
Before	1.05%	2.39%	1.72%	0.31%	0.24%	2.37%	0.80%	0.60%
After	0.73%	1.41%	1.50%	0.69%	0.16%	0.64%	0.95%	0.00%
	Members							
	9	10	11	12	13	14	15	16
Before	0.73%	0.04%	1.32%	1.41%	3.19%	0.65%	0.77%	0.04%
After	0.93%	0.01%	1.68%	1.52%	0.73%	0.33%	0.34%	0.00%

Figure 3 provides an alternative illustration of participant performance, comparing "before" and "after" pair-wise numbers. Figure 3 shows reductions in most participants printing after applying the feedback method. Table 2 shows that the average reduction is approximately 28% compared to printing before feedback was received. Nevertheless, five participants, namely 4, 7, 9, 11, and 12, show slight increases in proportional printing.

In order to infer if the feedback method did in fact cause a reduction in the average proportional printing, we performed a statistical test over two samples: before and after applying the feedback method. These two samples are in fact repeated measures of the same participant printing. Thus, a standard statistical paired t-test is suitable in this case [23] to test our hypothesis. The t-test is done over one sample that represents

the pair-wise differences between measurements in the two samples. Because the "difference" sample size in this experiment is 16 < 30, the t-test cannot be done unless the "difference" sample follows a normal distribution.

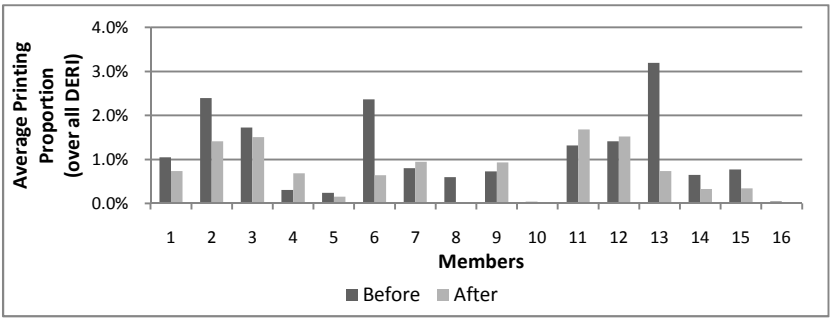


Fig. 3. Participant printing performance before and after applying the feedback method

In order to check normality of the "difference" sample, a standard Anderson-Darling test [24] with a standard level of confidence (commonly used in inferential statistics is $\alpha = 0.05$ [25]) was done. The Anderson-Darling test hypothesizes that the sample follows a normal distribution and tries to reject this hypothesis. In our case, the test results in a *P*-value = 0.343 which is greater than $\alpha = 0.05$. Consequently, the null hypothesis cannot be rejected, and the "difference" sample is inferred to be normally distributed. Additionally, Figure 4 illustrates the Q-Q plot of the "difference" random variable. It plots the variable against the normal distribution and shows the measures falling randomly around the $y=x$ line meaning that the sample follows a normal distribution.

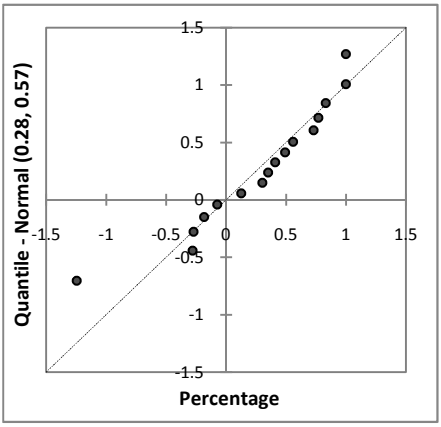


Fig. 4. Q-Q plot of the "difference" sample confirming a normal distribution

With the normality assumption checked, the paired *t*-test treats the hypothesis that the difference population mean is zero versus the hypothesis that it is less than zero.

Let us consider the standard level of confidence $\alpha = 0.05$. Applying a paired *t*-test to the values presented in Table 2 results in a *P*-value = 0.025. As the obtained *P*-value is below the 0.05 level of confidence, it is safe to reject the null hypothesis and conclude that the population mean of the "after" sample is less than the population mean of the "before" sample. Thus, it is statistically significant to say that the feedback method helped decrease the individuals' proportional printing after it was applied.

4.2 Team Level Printing Performance

Improving the teams printing performance is not the ultimate goal of the team-based feedback method. However, if the feedback is effective at persuading individuals within those teams, it is a natural consequence. In fact, it is possible to achieve statistically significant improvements at the individual level while failing at the team level if there are outliers who increase printing in a way that affects the sum of the overall team performance, but not the statistical significance of a large number of members who decrease their printing. Figure 5 illustrates the before and after team average proportional printing, measured as the average of sums of each team members proportional printing. All teams except unit 2 were able to reduce overall team printing.

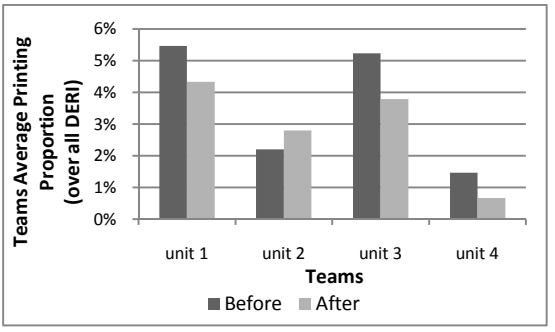


Fig. 5. Teams printing performance before and after applying the feedback method

5 Discussion

In order to analyse the results of our experiment, we base the discussion on previous work by Medland [2]. Medland conducted interviews with key stakeholders to examine the attitudes and feedback preferences for printing consumption data. His study revealed three themes which we re-examine in light of our own study:

Theme 1: any printing conducted by staff was viewed as necessary, excepting accidents.
Theme 2: staff viewed different metrics as relevant or alternatively, as unnecessary for feedback on printing behaviour.
Theme 3: staff viewed co-workers championing reduced paper use as a helpful reminder to be conscious of how much they were printing.

5.1 Theme 1: All Printing Is Necessary

As presented in Section 4.1, five participants – about 31% of the participants – showed a slight increase in their proportional printing and thus did not respond to the feedback method. This observation suggests that an important incentive to persuade people to reduce their printing was not addressed by the team-based feedback method. This finding presents an interesting avenue for future qualitative research.

While the feedback method covers well themes 2 and 3 as discussed in Sections 5.2 and 5.3, it is likely that the driving force behind the non-responsive people not to reduce their printing is due to theme 1, i.e. their view that all their printing is rather necessary and cannot be avoided. Thus, this experiment raises attention to theme 1 found by Medland [2] and suggests that any feedback method needs to address fully or partially this theme to get better results, e.g. by changing business processes.

5.2 Theme 2: Flavoured Feedback

Theme 2 suggests that different people have different views of what types of feedback are valuable or useful to them, such as traditional or eco-metric. Emails sent by our feedback software in this experiment applied these feedback methods in tandem. The empirical study showed a statistically significant proof that the feedback method helped decrease the printed-paper consumption in combination. Thus, stating the same information in different ways that address different models of individual perception is in fact a good feature of feedback methods confirming ideas presented in [2].

5.3 Theme 3: Comparison and Competition

Emails sent by the feedback software included two types of comparisons: temporal and social. Temporal comparison contrasts team printing performance in a week to performance the week prior. Social comparison compares printing performance on a team-to-team basis. This study evaluated both types on a team-level rather than on an individual level. Results showed that the feedback method applied is sound and useful, as it has been able to form an incentive for individuals to lower unnecessary printing. That is consistent with theme 3 presented previously in [2].

Nevertheless, our experiment showed that the overall printing performance of some teams may not change even when the majority of team members are responsive. We posit that this is due to the fact that outliers may exist in a team with the feedback method unable to motivate them to reduce their usage. This raises interest in hybrid individual and team-based feedback models where predefined outliers can be targeted with personal feedback emails and other people with team-based feedback emails.

6 Conclusions and Future Work

This paper reported an empirical study conducted in the Digital Enterprise Research Institute (DERI) over a year from October 2010 to November 2011. This study

examined the hypothesis that a team-based feedback method can reduce individual consumption of a resource such as sheets of paper. Empirical results confirmed the hypothesis and showed an overall average reduction in paper usage of about 28% among the participants in DERI.

Of continuing interest to the researchers is the understandable function of a follow-up qualitative study, investigating participant experience, change of printing habits, and team dynamics. Qualitative methods would service a review of our already published findings, and provide rich in-depth accounts from individuals and teams, helping to strengthen and extend our findings in numerous ways. Discovering whether elements such as competition or communities of practise emerge is a research avenue of real value.

Future quantitative work would consider testing out the hypothesis with different types of resources such as electricity or non-renewable energy. Now that we have a significant result we would also seek to introduce further experimental conditions to strengthen our conclusions, such as providing another team with only verbal announcements "please print less", or irregular and less granular feedback. A possible direction also is to leverage different forms and definitions of teams that go beyond top-down teams defined by organisational charts. An example can be bottom-up teams formed by participants in a work place dedicated for improving resource consumption. Another example is virtual online communities such as those formed on social networks.

Acknowledgements. This work has been funded by Science Foundation Ireland under Grant No. SFI/08/CE/I1380 (Lion-2), and co-sponsored by a Queensland Government Smart Futures Fellowship and National ICT Australia (NICTA). NICTA is funded by the Australian Government as represented by the Department of Broadband, Communications and the Digital Economy and the Australian Research Council through the ICT Centre of Excellence program.

References

1. Pierce, J., Odom, W., Blevis, E.: Energy aware dwelling: a critical survey of interaction design for eco-visualizations. In: Proceedings of the 20th Australasian Conference on Computer-Human Interaction: Designing for Habitus and Habitat, pp. 1–8. ACM, New York (2008)
2. Medland, R.: Curbing paper wastage using flavoured feedback. In: Proceedings of the 22nd Conference of the Computer-Human Interaction Special Interest Group of Australia on Computer-Human Interaction, pp. 224–227. ACM, New York (2010)
3. Fogg, B.J.: Creating persuasive technologies: an eight-step design process. In: Proceedings of the 4th International Conference on Persuasive Technology, pp. 44:1–44:6. ACM, New York (2009)
4. Fogg, B.J.: Persuasive technology: using computers to change what we think and do. Ubiquity 2002 (2002)
5. Lockton, D., Harrison, D., Stanton, N.A.: The Design with Intent Method: A design tool for influencing user behaviour. Applied Ergonomics 41, 382–392 (2010)

6. Khaled, R., Biddle, R., Noble, J., Barr, P., Fischer, R.: Persuasive interaction for collectivist cultures. In: Proceedings of the 7th Australasian User Interface Conference, vol. 50, pp. 73–80 (2006)
7. Foth, M., Paulos, E., Satchell, C., Dourish, P.: Pervasive computing and environmental sustainability: two conference workshops. IEEE Pervasive Computing 8, 78–81 (2009)
8. Davis, F.D.: Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Quarterly, 319–340 (1989)
9. Kankanhalli, A., Tan, B.C.Y., Wei, K.K.: Contributing knowledge to electronic knowledge repositories: An empirical investigation. Mis Quarterly, 113–143 (2005)
10. Reno, R.R., Cialdini, R.B., Kallgren, C.A.: The transsituational influence of social norms. Journal of Personality and Social Psychology 64, 104 (1993)
11. Schultz, P.W., Nolan, J.M., Cialdini, R.B., Goldstein, N.J., Griskevicius, V.: The constructive, destructive, and reconstructive power of social norms. Psychological Science 18, 429–434 (2007)
12. Cialdini, R.B.: Crafting normative messages to protect the environment. Current Directions in Psychological Science 12, 105–109 (2003)
13. Darby, S.: The effectiveness of feedback on energy consumption. A Review for DEFRA of the Literature on Metering, Billing and direct Displays, 486 (2006)
14. 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 (2010)
15. Strengers, Y.A.A.: Designing eco-feedback systems for everyday life. In: Proceedings of the 2011 Annual Conference on Human Factors in Computing Systems, pp. 2135–2144 (2011)
16. Lee, T.B., Hendler, J., Lassila, O.: The semantic web. Scientific American 284, 34–43 (2001)
17. Klyne, G., Carroll, J.J.: Resource Description Framework (RDF): Concepts and Abstract Syntax, `http://www.w3.org/TR/2004/REC-rdf-concepts-20040210/` (last accessed January 2013)
18. Berners-Lee, T.: Linked Data- Design Issues, `http://www.w3.org/DesignIssues/LinkedData.html` (last accessed January 2013)
19. Curry, E., Hasan, S., ul Hassan, U., Herstand, M., O'Riain, S.: An Entity-Centric Approach to Green Information Systems. In: The 19th European Conference on Information Systems, ECIS (2011)
20. Prud'Hommeaux, E., Seaborne, A.: SPARQL query language for RDF. W3C working draft. 4 (2008)
21. Curry, E., Hasan, S., O'Riain, S.: Enterprise Energy Management using a Linked Dataspace for Energy Intelligence. In: Second IFIP Conference on Sustainable Internet and ICT for Sustainability (2012)
22. Curry, E., Hasan, S., White, M., Melvin, H.: An Environmental Chargeback for Data Center and Cloud Computing Consumers. In: Huusko, J., de Meer, H., Klingert, S., Somov, A. (eds.) E^2DC 2012. LNCS, vol. 7396, pp. 117–128. Springer, Heidelberg (2012)
23. Hsu, H., Lachenbruch, P.A.: Paired t test. Wiley Encyclopedia of Clinical Trials (2008)
24. Anderson, T.W., Darling, D.A.: A test of goodness of fit. Journal of the American Statistical Association 49, 765–769 (1954)
25. Cowles, M., Davis, C.: On the origins of the. 05 level of statistical significance. American Psychologist 37, 553 (1982)

Three Approaches to Ethical Considerations in the Design of Behavior Change Support Systems

Pasi Karppinen and Harri Oinas-Kukkonen

Department of Information Processing Science, University of Oulu, Finland
{pasi.karppinen,harri.oinas-kukkonen}@oulu.fi

Abstract. Many ethical questions arise when developing persuasive systems. It has become evident that there is no silver bullet which would make it easy to resolve all ethical issues in all cases. This paper seeks to analyze and define potential ways to address ethical considerations in persuasive systems design. We suggest that there are three main approaches: a guideline-based approach, stakeholder analysis, and involving users. This paper helps to understand the strengths and weaknesses of these approaches when developing behavior change support systems, which, by their very nature, request deep engagement and commitment from their users. A pragmatic goal for this paper is to help designers choose an approach for their projects at hand.

1 Introduction

Information technology is never neutral [1]. It influences people's attitudes and behaviors in one way or another, and its developers have to be aware of the full power it exercises over its users [2]. Persuasive and behavior change support systems (BCSS) are designed to create a cognitive and/or emotional change in the mental state of a user [2]. However, due to the many challenges involved, attempting to change users' behaviors or attitudes through these systems can become something of an ethical minefield [3].

The studies of Berdichevsky and Neuenschwander [4] and Fogg [5], [6] were the first academic works to directly address the ethical issues concerning persuasive technology and design. More recently, there has been a growing interest in persuasive ethics, as evidenced by the studies of Burri Gram-Hansen [7], Davis [3], Smids [8], Spahn [9], and Yetim [10]. These studies are convincing pieces of work, but their suggestions for resolving ethical issues differ from one another. Whereas Berdichevsky and Neuenschwander [4] present eight moral principles for the designer to follow, Davis [3] and Yetim [10] argue that designers should find consensus of ethical issues with the stakeholders. Smids [8] concludes that the most important moral question for persuasive technology is the person's voluntariness of change, whereas Spahn [9] sees persuasion as an act of communication, which should always follow validity claims of speech acts. Fogg [5] suggests stakeholder analysis for examining the ethics of complicated situations, whereas Burri Gram-Hansen [7] introduces ethics as an intuitive result of human nature, rather than as a reason-based rule.

S. Berkovsky and J. Freyne (Eds.): PERSUASIVE 2013, LNCS 7822, pp. 87–98, 2013.

It is hard to argue that any of the abovementioned studies would solely address and help resolve all possible ethical considerations. It is much more fruitful to treat them as potential ways to address different types of ethical challenges. A designer can, of course, build upon lessons learned and academic studies, but it remains open as to how to systematically chooses a suitable ethical approach. The research question for this paper can perhaps be best phrased as: How can we model current approaches in the field of persuasive technology, which address ethical considerations when developing BCSS?

Thus, the goal of this paper is to aid in addressing the ethical considerations in designing persuasive and behavior change support systems. We emphasize the designs of BCSS, as they request deep engagement and commitment from their users, and, thus, many of the ethical questions in them require immediate concern. We propose a conceptual framework, which helps designers choose an approach for the design work at hand; therefore, the contribution of this paper is both pragmatic and academic. We recognize the major categories for approaching ethical considerations, and analyze eight related studies through the framework.

Section 2 will introduce how principles related to BCSS evoke ethical issues that may not be familiar to all persuasive systems. Section 3 concentrates on ethical issues recognized by the Persuasive Systems Design model developed by Oinas-Kukkonen and Harjumaa [1], [11]. In section 4, the framework is introduced. Section 5 discusses how the findings contribute to the actual design work, and examines the strengths and weaknesses of the different approaches.

2 Ethical Considerations Recognized by the BCSS Approach

Persuasive technology can be seen as a field of research, whereas a BCSS is an object of study [12]. Oinas-Kukkonen [2] defines BCSS as follows:

> A behavior change support system (BCSS) is a socio-technical information system with psychological and behavioral outcomes designed to form, alter or reinforce attitudes, behaviors or an act of complying without using coercion or deception.

As mentioned previously, Smids [8] states that the most important ethical question regarding persuasive technology is the user's voluntariness of change. Smids [8] builds his argument on the demarcation of persuasion, coercion, and manipulation. If BCSS turns out not to rely on voluntary change, it should not be called, sold, or promoted as a persuasive system [8]. BCSS emphasizes autogenous approaches, where a person uses technology to change his/her own behavior or attitude to his/her own goal [2]. Applications that promote healthier lifestyles are typical examples of BCSS. Spahn [9] points out that in self-persuasion, a user already shares the value in question, and uses the persuasive technology only to overcome a 'weakness of the will'. He argues that these cases are morally less problematic than using technology to

persuade others [9]. Even though the importance of voluntariness seems self-evident, it is not without contradictions, as we will address later.

There are two important steps for measuring behavioral changes: analysis of the persuasion context and analysis of the persuasive potential of the system [2]. For both methods, the O/C matrix is used as a means to analyze the intent of a BCSS, and the PSD model is used as a means to analyze the persuasive potential of the system, unveiling a great part of what BCSS are all about [2].

The O/C matrix developed by Oinas-Kukkonen [12] helps to analyze the intent and the outcome of a persuasive system. Successful outcomes in the matrix are the formation, alteration, or reinforcement of attitudes, behaviors, or compliance. A forming outcome (F-Outcome) stands for the birth of a pattern for a situation that did not previously exist. In practice, stopping a behavior also results in a new behavior (F-Outcome). An altering outcome (A-Outcome) stands for changes in a user's response to an issue, like increasing exercise. A reinforcing outcome (R-Outcome) stands for the reinforcement of current behaviors or attitudes, which makes them more resistant to change.

In the O/C matrix, the changes are also divided into three categories: a change in the act of complying, a behavior change, or an attitude change (C-, B-, and A-Change, respectively). The goal of the C-Change is simply to make sure that the person complies with the system's requests. The goal of a healthcare application can, for instance, guarantee that its user takes his/her daily medication. A system supporting a B-Change aims to elicit a deeper behavior change, rather than mere compliance. A one-off behavior change is naturally easier to achieve, whereas a long-term behavior change is clearly more difficult. The goal of A-Change is to influence a person's attitudes, rather than just their behavior. According to Oinas-Kukkonen [12], change-in-full occurs only when attitude change takes place, and a sustainable B-Change happens only through an A-Change.

As reinforcing outcome implies, behavior change does not need to be the transformation from one position to another. The importance of reinforcement is evident when treating addictions, such as smoking or abuse of alcohol. A person 'technically' quits their previous behavior only once, but he/she needs support to continue with the new behavior. Reinforcement is vital, and even though it is no longer a concrete, measurable change in behaviors, support and encouragement are required to keep the end-user's feet on the right path. For another related matter, BCSS can be built on a therapy routine that requires a great deal of compliance from the user. Compliance is not coercion, but, admittedly, it does not fully resonate with the principle of voluntariness either. Spahn [9] sees that technology can be regarded as the implementation of a technological paternalism, which conflicts with the ideal of the free and autonomous choice of the individual. When complying, the person may not have the proper motivation for doing so, and the key becomes to provide triggers for the user to take action and to comply with the requests of the application [2]. Oinas-Kukkonen [2] argues that there are numerous computer applications that use the same design principles and techniques as systems supporting C-Change, and that most research in the area of behavioral change actually focuses on C-Change.

3 Ethical Considerations Recognized by the PSD Model

The Persuasive Systems Design model (PSD) is a vehicle for designing and evaluating BCSS [2]. The PSD model includes the analyses of intent, event, and strategy of persuasion, and also discerns opportune moments for delivering the message(s) [2]. The PSD model defines software features for BCSS that are divided to four categories: primary task support, computer–human dialogue support, perceived system credibility, and social influence [11], [1]. Before analyzing the context or considering persuasive features, the designer should obtain a deeper understanding of persuasion on a postulate level. The seven postulates common to all BCSS [2, cf. 1] are:

(P1) information technology is never neutral, but rather it always influences its user(s) in one way or another;
(P2) people like their views to be organized and consistent;
(P3) persuasion is often incremental;
(P4) the direct and indirect routes are key persuasion strategies;
(P5) BCSS should be both useful and easy to use;
(P6) persuasion through BCSS must always be unobtrusive to a user's primary tasks; and
(P7) persuasion through BCSS should always be transparent.

These postulates have been defined based on both software design and psychology. For instance, as postulate P5 argues, if a system is useless or difficult to use, it is most probably hardly persuasive at all [2]. In this paper, we will focus on postulates regarding the route (P4) and transparency (P7), as they have very strong implications for the voluntariness principle.

Postulate P4 derives from theories such as the Elaboration Likelihood Model (ELM), and claims that using an indirect route in persuasion is one of the key persuasion strategies [13]. According to ELM, the central route is more enduring, resistant, and predictive of behavior [13], but since people are often too busy with their everyday lives, the indirect route becomes sometimes the only alternative. But, as Spahn [9] argues, persuasion should be based on prior consent. For a person to do things voluntarily, he/she must have clear picture of what he/she is agreeing on [9]. According to Oinas-Kukkonen [2], there may actually also be situations where computer-mediated persuasion takes place without the user being aware of it. Changing user behavior with the help of subliminal triggers greatly challenges the voluntariness concept. Smids [8] argues that 'unconscious persuasion' is an oxymoron that violates the voluntariness condition, but that lack of awareness on the other hand does not necessarily make persuasive technology manipulative [8], which leaves room for indirect persuasion to be considered ethical as well.

Postulate P7 requires persuasion to be transparent, and underlines the need for revealing the designer's bias behind a BCSS [2]. Atkinson [14] and Davis [3] argue that a user must be informed of the persuaders' intent for persuasion to be ethical. Transparency supports also voluntariness, because it is the only way for users to evaluate if the persuasive system is ethical based on their own moral standards. Berdichevsky

and Neuenschwander's statement [4], however, reveals one problem concerning transparency:

> The creators of a persuasive technology should disclose their motivations, methods, and intended outcomes, except when such disclosure would significantly undermine an otherwise ethical goal.

On the one hand, without transparency users have no choice but to rely on the moral choices the designer has made. On the other hand, if the designer is confident of his/her goals being ethical, why pursue the need for transparency? To some extent, when including the exception for the rule, Berdichevsky and Neuenschwander [4] weaken the role of disclosure, but then also point to the fact that making the methods of persuasion fully transparent may undermine the outcome. Thus, level of openness remains inconclusive in their definition.

As seen above, persuasion is a tricky concept to apply ethical principles to. Atkinson [14] asks that if the user has voluntarily chosen to use a system to achieve a predetermined goal, can we talk about persuasion at all? At the same time, Atkinson [14] refers to Hart's natural right theory, and claims that freedom is an inalienable fundamental human right. This can appear paradoxical at first, but there is a fruitful thought that links together with BCSS. According to Atkinson [14], ethical safeguarding can be achieved if the purpose of the persuasion is exposed at the beginning of one's engagement with a system. Hence, a user can choose whether to accept or reject the BCSS's offering. We see that voluntariness and transparency are important values for persuasive technology, and that the postulates P4 and P7 play a big role when a person is engaging with the system. However, we claim that all persuasive acts during the change process do not need to be voluntary or fully transparent.

Oinas-Kukkonen [2] makes the distinction between BCSS and persuasive systems as follows:

> A BCSS places more emphasis on the actual outcome than a persuasive system, which, even if its developers were interested in the outcomes as well, in most cases emphasizes more the persuader's intent than measuring the actual outcome.

Behavior change is often a long lasting process that requires commitment and compliance from the user, thus by considering merely the issues of voluntariness and transparency does not mean that the system is automatically on solid ethical ground. In other words, after the user has committed to use the system, this does not give the designer license to do whatever he/she pleases. Some of the BCSS design issues are naturally technical, but many are user-related and social; some may relate to organizations, cultures, or the whole of society [2]. There are a growing number of BCSS that are built jointly with other stakeholders, including end-users [15]. Some aim to improve the wellbeing of entire community. In sum, we see that when designing a BCSS, users and other stakeholders must be taken into account also from an ethical perspective.

4 Ethical Framework

We propose here a conceptual ethical framework, which composes of three categories: guideline-based approaches, stakeholder analysis, and user involvement. The guideline-based category includes ethical approaches that provide general ethical precepts, but do not give explicit guidance on the consideration of users' and other stakeholders' moral norms. The stakeholder analysis category targets ethical approaches that evaluate stakeholders, without actually involving them in the design or the use process. The user involvement category includes approaches that take stakeholders' voices into account in order to seek ethical solutions with them.

We identified seven studies in the persuasive technology research field that focus on ethical issues: Berdichevsky and Neuenschwander [4]; Burri Gram-Hansen [7]; Davis [3]; Fogg [5]; Smids [8]; Spahn [9]; and Yetim [10]. We mapped these works to our ethical framework, and later decided to add Friedman et al.'s [16] study to place more emphasis on the stakeholder analysis category, even though the study does not discuss the persuasive technology field as such. In addition, Davis [3] and Yetim [17] use Friedman et al.'s [16] ideas as a phase for involving users. The abovementioned studies are summarized in Table 1 with their corresponding category and primary ethical contribution from the BCSS design point-of-view.

The guideline-based category contains various approaches, which suggest general principles for addressing ethical issues. As seen also from Table 1, most research interest thus far seems to fall into this category. For example Berdichevsky and Neuenschwander [4] suggest eight principles for persuasive technology and design, which can be summarized as follows: 1) outcomes of persuasion must be considered ethical even without persuasion, or if technology were not involved; 2) motives of the persuader should not be considered unethical when using traditional ways of persuasion; 3) designers should take responsibility for all reasonably predictable outcomes of the technology's use; 4) designers must respect the users' privacy as they would respect their own; 5) sharing user's personal information to third parties should be scrutinized for privacy concerns; 6) designers must disclose their motivations, methods, and intended outcomes, except when such disclosure would significantly undermine an otherwise ethical goal; 7) persuasive technology should not misinform the user; and 8) designers should never seek to persuade a user to do something they would not wish to be persuaded to do themselves. The most important guideline is the eighth principle, "the Golden Rule of Persuasion" [4]. Burri Gram-Hansen [7] arrived at the same conclusion, and states that we must strive towards doing to others as we hope others will do unto us. A designer can use the golden rule as a main principle to judge all actions and determine whether any possible action is considered ethical.

The guidelines suggested by Berdichevsky and Neuenschwander [4] are in active use, such as Kaptein and Eckles' [18] recent study on ethical considerations related to persuasion profiling. Berdichevsky and Neuenschwander [4] also specify and emphasize the value of privacy in their guideline, however Kujala and Väänänen-Vainio-Mattila [19] show that values can be retrieved from various different theoretical frameworks, and privacy is just one of many principles. Whether we like it or not,

Table 1. Framework of ethical approaches in persuasive technology design

Approach	Publication	Primary ethical contribution for BCSS designer
Guideline	Berdichevsky & Neuenschwander [4]	Eight principles for persuasive technology design, from which the golden rule is considered most important: The creators of a persuasive technology should never seek to persuade a person or persons of something they themselves would not consent to be persuaded to do.
	Burri Gram-Hansen [7]	Ethical reflections are intuitive and personal. Strive to create a product that will have impact on the user in a way which yourself accept as ethically acceptable.
	Smids [8]	The most important ethical question regarding persuasive technology is the person's voluntary desire for change. Do not use techniques of coercion, manipulation, or subliminal persuasion.
	Spahn [9]	Three principles for persuasion: 1) persuasion should be based on prior consent; 2) ideally the aim of persuasion should be the end of the persuasion; 3) persuasion should grant as much autonomy as possible to the user.
Stakeholder analysis	Fogg [5]	Seven step stakeholder analysis: 1) list all stakeholders; 2) list what each stakeholder can gain and 3) what they can lose; 4) evaluate which stakeholder has most to gain and 5) most to lose; 6) determine ethics by examining gains and losses in terms of values; 7) acknowledge your own values that you bring to the analysis.
	Friedman et al. [16]	Values are retrieved from stakholders through analysis that has three different layers: conceptual, empirical and technical investigations.
User involvement	Davis [3]	Participation in design. Usage of value-sensitive design to evaluate the values of the direct and indirect stakeholders, and participatory design to involve potential users as full participants in the design process.
	Yetim [10]	Conversation in use time. A total of 21 critical questions that guide reflections on systems for three discourse types: pragmatic (goal-value, action-goal, action-value); ethical (identifying, checking); and moral (identifying, checking).

guideline-based approaches are fundamentally subjective. What a designer might consider to be ethically correct may be entirely unethical from a user's perspective.

The stakeholder analysis approach aims to evaluate the values of different stakeholders, and it holds the idea that values vary from one situation to the next, and that there are no easy answers that would satisfy all [5]. The approach is widely used in business ethics [20], [21], and stems from when the traditional shareholder view was

considered too narrow and ethically insufficient [21]. The stakeholder analysis also aims to consider those groups which do not have the power to make decisions or to participate in decision-making processes, but are nonetheless dependent of the decisions being made [21]. Stakeholder analysis does not mean that suggested ethical guidelines have no meaning. For instance, Fleischman and Wallace [22] argue that transparency is an essential tool to empower and preserve users' autonomy, also from the multi-actor viewpoint.

In the persuasive technology field, Fogg [5] identifies seven steps for stakeholder analysis (see Table 1), and argues that the key is for persuasive technology designers to become aware of the range of ethical issues involved. The Value Sensitive Design (VSD) methodology by Friedman, Kahn and Borning [16] is considered the most comprehensive framework for advancing value-centered research [17]. They introduce ten steps for conducting VSD, wherein the key element is the stakeholder analysis [16]. According to Davis [3], the approach emphasizes values of moral import, and thus speaks to ethical concerns surrounding technology design; it also reveals situations in which designers must make tradeoffs between conflicting values, perhaps based on interviews to collect stakeholders' values [16]. Even if stakeholder analysis by no means implies that the designer should remain isolated behind the office desk, it does not really involve users in the design process.

When stakeholder analysis considers stakeholders as the object of study, user involvement approaches seek ethical solutions in terms of these stakeholders. Davis [3] and Yetim [10] have carried out compelling studies on the suggested framework's user involvement category. In fact, their approaches are so different it would be possible to split this category into two streams: participation in design and conversation in use time. Davis [3] uses the VSD [16] method to identify the values of both direct and indirect stakeholders, and takes one step further to engage stakeholders as equal participants in design. When Davis [3] leans on the principles of participatory design, she aims to draw more attention to the welfare of indirect stakeholders, as these people not only use the technology, but are also affected by its use. In her later study, Davis [23] introduces the Inspirational Card Workshop as a method to address ethical issues in participatory design.

To address multiple voices of different stakeholders, Yetim [10] uses Jürgen Habermas' discourse ethics, where the decisive power is given to the most compelling argument. He also underlines that the method is valid also in terms of use time, and not merely in design [17], [10]. In discourse ethics, every stakeholder must be equal in order to participate in the discussion, and instead of settling for a midpoint compromise the agreement should be based on the jointly agreed best argument [24]. Every stakeholder's voice is to be taken seriously, and since the decisive power is given to the most compelling argument [24], Yetim [10] proposes in his study a hierarchy of critical questions that helps identify and decide on ethical arguments during the debate.

Thus far, few published studies on persuasive technology deal with ethical issues. New ethical approaches will be published eventually, and we believe that scholars and designers will be able to map them accordingly to the presented categories in our framework.

5 Discussion

We began this study by claiming that information technology is never neutral. It is indeed constantly affecting our lives. The internet, for example, continues to change businesses, software design, the way we perceive people, and the skills required of us [25]. More specifically, this paper presented a framework for recognizing and choosing suitable ethical approaches for persuasive systems design tasks at hand. In this, voluntariness and transparency were found to be important values.

To make a system as ethically safe as possible along the way, it seems tempting for designers to involve stakeholders in the design process to build ethical consensus through participation and/or conversation. Ethical approaches in the user involvement category are especially useful when the designed system is targeted at a specific and predefined group, i.e. when the goal of conversation can be specific enough, or when building a system jointly with equal partners. However, designers often have only limited resources to do such thorough work in reality.

The user involvement category has two handicaps that designers must take into account. Firstly, building consensus on key values does not make the system automatically ethical. Even though moral values should be emphasized, other values should not be neglected [16]. For instance, monetary wealth is a value too, and some stakeholders may be very keen to have it as a primary goal for the system, whereas others would hold the totally opposite view. Secondly, identifying in addition to direct also indirect stakeholders can be an extremely challenging task. Local solutions can be considered in a broader context, and they can be lifted even to a global scale when thinking about the consumption of natural resources, for instance. The question is when and where to stop the identifying process. Thirdly, designers should also keep in mind that users in most cases do not explicitly think about their values, and can have problems to articulate them, especially when these values are unconscious or perhaps socially unacceptable.

Reaching consensus itself is often hard, and in large software development projects this may be an overly optimistic goal. Developers in responsible positions should be prepared to solve stalemates in such a way that the solution remains ethically solid. Yetim's [10] hierarchies can be fruitful for these types of situations. Mingers and Walsham [24] argue that even the famous open source development project of Linux kernel cannot meet the ideal conditions of discourse ethics, when the inner core of developers have strong rights to control the implementation of changes. This means that contributors are not equally participating in the debate. This type of situation is very common in information systems development practice, where developers are the gatekeepers who make the final decisions on the technical solutions implemented.

Whether the design project is big or small, in order to design BCSS properly, users must be taken into account in multiple stages of design. Practically all systems are built for some target group in mind, and it is only practical to make ethical considerations along the way. If the designer has difficulties imagining himself/herself in someone else's shoes, he/she can interview stakeholders who will be involved with the system. Stakeholder analysis does not demand an overwhelming amount of work, and there are methods such as the Value Senstitive Design that help to identify users'

values; Kujala and Väänänen-Vainio-Mattila even describe the potential varieties of values, which makes identifying values easier [19]. Stakeholder analysis seems to be a particularly suitable approach for situations where the use context of the system [cf. 15] is clear, as value-based product solutions often are highly context-dependent [19]. Stakeholder analysis is suitable also to identify values, but similarly to user involvements it says nothing about the norms stakeholders should follow [21]. It seems to be that in the field of persuasive technology, the stakeholder analysis category is yet to receive more of the rigorous scientific investigation. Nevertheless, it may turn out to be a very useful method to aid design.

In the current business climate, it is very easy to launch a BCSS to global markets, which makes the identification of stakeholders and analyzing their values ever more important, despite the fact that this is a time-consuming and wearisome task. Guidelines may operate as checklists to tackle important ethical issues. As a definition, BCSS do not deceive, manipulate, or coerce users; they should be transparent to enable an individual's free choice to engage with the system [2]. Users rely on BCSS developers as experts of their topic, and expect the system to deliver what it promises. However, if the system is badly designed and a user does not have a chance to achieve his/her goal for this reason, it can be considered unethical as well. Professionals often regard ethics as a source of edification, and fail to realize that they encounter ethical issues and challenges in the ordinary course of their work [26]. However, questions of appropriate actions can arise partly in situations where nobody has done anything clearly wrong [26].

In overall, BCSS are deliberately designed to change users' behavior, and, as Berdichevsky and Neuenschwander stated [4], we as designers ought to take responsibility for the outcomes of a system's use that are reasonably predictable. However, it remains practically impossible for designers to predict all outcomes for all stakeholders, based on their own limited perspectives [3], and no experimental approaches have the ability to solve all questions of morality and ethics when designing BCSS. Thus, the designers should choose a suitable approach for the task at hand based on an awareness of their values and the values which will be brought into the design. Better yet, values should be explicitly specified and exemplified. At times, a designer faces a situation where he/she has no other option, but to simply count on his/her own reasoning or perhaps even intuition. Being faithful to one's intuition is not naïve nonsense [cf. 7]. If you feel that what you are doing is unethical – just do not do it.

6 Conclusion

The motivation for this study rises from the important goal to support ethical design of persuasive and behavior change support systems. Due to the lack of a single approach that could be used in all design cases, designers need conceptual help to choose a suitable approach for the design case at hand. By using the BCSS design principles [2] and the PSD design model [1], it was possible to recognize the major ethical approaches, differentiate how they deal with issues related to users and other stakeholders, and propose a conceptual framework to help persuasive designers and

researchers. The ethical framework comprises three major categories: guideline-based approaches, stakeholder analysis, and user involvement. The framework is unique for persuasive technology and design, and a special characteristic is that it is built from the design perspective, rather than from philosophical traditions, focusing on the practical need to solve ethical questions in the design. Moreover, philosophically oriented researchers may benefit from this framework by being able to recognize areas in which designers urgently need ethical guidance. In the proposed framework, the values of voluntariness and transparency play a vital role for a user to be able to decide whether to commit to the use of a BCSS. However, even they, or other guidelines, do not guarantee a system's high ethical status.

In conclusion, the current base of knowledge is still very limited for addressing ethical considerations in the field of persuasive technology. There is a need for more studies regarding the ethical design and development of persuasive and BCSS. The evidence for the proposed framework is also limited. Thus, we invite designers and researchers to put the framework into practice, to scrutinize it under evaluation, and to report their lessons learned.

Acknowledgements. We wish to thank the anonymous reviewers of this paper for their insightful comments. The study was partially supported by the SalWe Research Program for Mind and Body, grant 1104/10, and the Someletti research project on Social Media in Public Space, grant 1362/31, both provided by Tekes, the Finnish Funding Agency for Technology and Innovation.

References

1. 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)
2. Oinas-Kukkonen, H.: A Foundation for the Study of Behavior Change Support Systems. Personal and Ubiquitous Computing, 1–13 (2012)
3. Davis, J.: Design Methods for Ethical Persuasive Computing. In: 4th International Conference on Persuasive Technology, pp. 1–8. ACM, NY (2009)
4. Berdichevsky, D., Neuenschwander, E.: Toward an Ethics of Persuasive Technology. Communications of the ACM 42, 51–58 (1999)
5. Fogg, B.J.: Persuasive technology: Using computers to change what we think and do. Morgan Kaufman, San Francisco (2003)
6. Fogg, B.J.: Persuasive Computers: Perspectives and Research Directions. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 225–232. ACM, NY (1998)
7. Gram-Hansen, S.B.: Towards an Approach to Ethics and HCI Development Based on Løgstrup's Ideas. In: Gross, T., Gulliksen, J., Kotzé, P., Oestreicher, L., Palanque, P., Prates, R.O., Winckler, M. (eds.) INTERACT 2009. LNCS, vol. 5726, pp. 200–203. Springer, Heidelberg (2009)
8. Smids, J.: The Voluntariness of Persuasive Technology. In: Bang, M., Ragnemalm, E.L. (eds.) PERSUASIVE 2012. LNCS, vol. 7284, pp. 123–132. Springer, Heidelberg (2012)

9. Spahn, A.: And Lead Us (Not) into Persuasion...? Persuasive Technology and the Ethics of Communication. Science and Engineering Ethics 18, 1–18 (2011)
10. Yetim, F.: A Set of Critical Heuristics for Value Sensitive Designers and Users of Persuasive Systems. In: ECIS 2011 Proceedings, Helsinki (2011)
11. 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)
12. 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)
13. Petty, R.E., Cacioppo, J.T.: Communication and Persuasion: Central and Peripheral Routes to Attitude Change. Springer, New York (1986)
14. Atkinson, B.M.C.: Captology: A Critical Review. In: IJsselsteijn, W.A., de Kort, Y.A.W., Midden, C., Eggen, B., van den Hoven, E. (eds.) PERSUASIVE 2006. LNCS, vol. 3962, pp. 171–182. Springer, Heidelberg (2006)
15. Ahola, H., Oinas-Kukkonen, H., Koivumaki, T.: Customer Delivered Value in a Web-Based Supermarket. In: Proceedings of the 33rd Hawaii International Conference on Systems Sciences, HICSS 2000. IEEE Computer Society Press (2000)
16. Friedman, B., Kahn Jr., P.H., Borning, A.: Value Sensitive Design and Information Systems. In: Zhang, P., Galletta, D. (eds.) Human-Computer Interaction and Management Information Systems: Foundations, pp. 348–372. M.E. Sharpe, New York (2006)
17. Yetim, F.: Bringing Discourse Ethics to Value Sensitive Design: Pathways to Toward a Deliberative Future. AIS Transactions on Human-Computer Interaction 3, 133–155 (2011)
18. 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)
19. Kujala, S., Väänänen-Vainio-Mattila, K.: Value of Information Systems and Products: Understanding the Users' Perspective and Values. Journal of Information Technology Theory and Application 9, 23–39 (2009)
20. Goodpaster, K.E.: Business Ethics and Stakeholder Analysis. Business Ethics Quarterly 1, 53–73 (1991)
21. Stahl, B.C.: Ethical Issues of Information and Business. In: Himma, K.E., Tavani, H.T. (eds.) The Handbook of Information and Computer Ethics, pp. 311–359. John Wiley & Sons, Hoboken (2008)
22. Fleischmann, K.R., Wallace, W.A.: A Covenant with Transparency: Opening the Black Box of Models. Communications of the ACM 48, 93–97 (2005)
23. Davis, J.: Generating Directions for Persuasive Technology Design with the Inspiration Card Workshop. In: Ploug, T., Hasle, P., Oinas-Kukkonen, H. (eds.) PERSUASIVE 2010. LNCS, vol. 6137, pp. 262–273. Springer, Heidelberg (2010)
24. Mingers, J., Walsham, G.: Toward Ethical Information Systems: The Contribution of Discourse Ethics. MIS Quarterly 34, 833–854 (2010)
25. Oinas-Kukkonen, H., Oinas-Kukkonen, H.: Humanizing the web: Change and social innovation. Palmgrave Macmillan, Basingstoke (2013)
26. Goodman, K.W.: Health Information Technology: Challenges in Ethics, Science, and Uncertainty. In: Himma, K.E., Tavani, H.T. (eds.) The Hand-book of Information and Computer Ethics, pp. 293–309. John Wiley & Sons, Hoboken (2008)

Using Log-Data as a Starting Point to Make eHealth More Persuasive

Saskia M. Kelders and Julia E.W.C. (Lisette) van Gemert-Pijnen

University of Twente, Department of Health, Psychology and Technology
{s.m.kelders,j.vangemert-pijnen}@uwente.nl

Abstract. Despite the large number of eHealth projects to date and the positive outcomes of evaluation studies, the adherence to eHealth interventions is lower than expected. To understand how persuasive technology can influence the adherence to eHealth interventions process data (log-data) about the usage of technology (system and content) can provide a starting point for employment of persuasive features into the design of technology. The log-data of the usage of an eMental health intervention used as an example in this paper, contained a record of actions taken by each participant with for each action the following information: user-id; action type; action specification; time and day. The log-data showed critical episodes for employment of persuasive components to increase adherence: episodes to determine the willingness to follow a therapy, awareness of their non-coping strategies, adoption of "new" skills for behavior change.

1 Introduction

All over the world eHealth is being increasingly introduced into the healthcare system to support access, safe-care and self-care management [1]. Despite the large number of eHealth projects to date and the positive outcomes of evaluation studies, the actual use of eHealth interventions is lower than expected [1-5]. Many projects fail to survive beyond the pilot phase and studies that investigate the effectiveness of eHealth applications most often do not show any long-term effects. Although many eHealth interventions reach a large group of participants, not all of these participants complete the intervention and might therefore not benefit as much from the intervention as they might. What's the problem?

In the field of health promotion, interventions are rather content-driven. Health-experts often translate a given content into a digital environment, in most cases Internet. Technology is not applied as an adaptive and interactive communication system that can be tailored and personalized to users and context of usage. Another related issue lies in the development process of eHealth interventions. Authors have advocated more user involvement and a more structured development process [2, 6-8]. However, many eHealth interventions are still developed in an ad-hoc manner, which may in part be due to the difficulties of developing in a multidisciplinary environment [2].

S. Berkovsky and J. Freyne (Eds.): PERSUASIVE 2013, LNCS 7822, pp. 99–109, 2013.

Often the development of the technology is expert or engineering driven and the development of technology and content is done separately instead of intertwined, which can lead to stand alone applications where there is no fit with users, content and system [2]. Finally, the capacities of technologies to monitor real-time usage are overlooked due to a merely outcome focus. Technologies have the power to be persuasive, to motivate people in a positive way and as such being a nudge or support for healthier living. In this paper we focus on the capacities of technology to monitor the usage and users of interventions aimed at healthier living. We show how continuous measurement of usage and users via logfiles can be the starting point for employment of persuasive triggers.

1.1 Log Files to Prompt Full Employment of Persuasive Technology

Logfiles can be used to get information about users and usage during real-time use of a technology based intervention. Based on the logfile data critical points can be identified for adherence to the intervention and for personalizing content and system to user's preferences.

Adherence to web-based interventions has been the subject of research for some time. Many studies focus on whether and which respondents' characteristics can explain variations in adherence [9]. Whether intervention or technology characteristics influence adherence has gained less attention, although there have been reviews that have explored this possibility [10,11]. These studies give insight into adherence as an outcome measure, but adherence can also be seen as a process. Adherence as a process measure can be described as the extent to which individuals experience the content of technology based interventions [12] and can be measured by observing the actual usage of the intervention (content & system) and the intended usage. Intended usage refers to the extent users should use the intervention. Given the purpose of lifestyle interventions an intended usage of an intervention is inherent to change behaviors, although in the description of interventions, intended usage is often not explained [12]. Measuring adherence should involve data on usage patterns, preferably on the level of the individual participant, because that will allow us to study how individuals interact with system and content and whether there are differences between adherers and non-adherers.

The adherence profile shows insight in usage and non-usage patterns and as such enables identification of critical points for employment of persuasive triggers. Also it is possible to get insight in differences in adherers and non-adherers with respect to the broad and in-depth usage of the intervention. Descriptive studies of freely accessible interventions have shown that they attract a considerable number of visitors, but that these visitors often interact with a fraction of the possible features of the intervention [13-19]. By observing the individual participants it is possible to determine what characterize users (user profile) and their preferences. This is important to personalize persuasive triggers to the motivations and capacities of a user for self-management.

1.2 The Show Case; "Living to the Full" Intervention

The show case for this paper is the web-based intervention "Living to the Full". This is an intervention for people with mild to moderate depressive symptoms to prevent the onset of a depression. The content of the intervention is based on ACT (Acceptance and Commitment Therapy) [20] and mindfulness [21,21] and has been published as self-help book [23]. The intervention has been shown to be effective in reducing depressive and anxiety symptoms as a group course and as a self-help course with email support [24-26]. The data used as an example in this paper comes from the web-based intervention that has been developed based on this content. The web-based intervention has been developed employing methods from the CeHRes Roadmap for eHealth development [2]. For in-depth information about the system and its performance see Kelders [12]. The web-based intervention included nine chronological lessons and each lesson consisted of psycho-educational material and exercises. When logging on to the web-based intervention, participants started in their 'cockpit' (Figure 1). From here, they could access all features of the intervention. Participants could access the web-based intervention at any time, from any place, free of charge. Only after finishing a lesson and receiving feedback, participants could proceed to the next lesson. This feedback was provided when a participant worked on the lesson for at least five days, viewed all psycho-educational material and completed all exercises. Participants were instructed to complete one lesson per week, but had twelve weeks in total to complete the nine lessons. Participants were free to choose whether they worked through a lesson in one session or in multiple sessions. It was estimated that participants would spend an average of three hours per week on the intervention (online and offline activities combined).

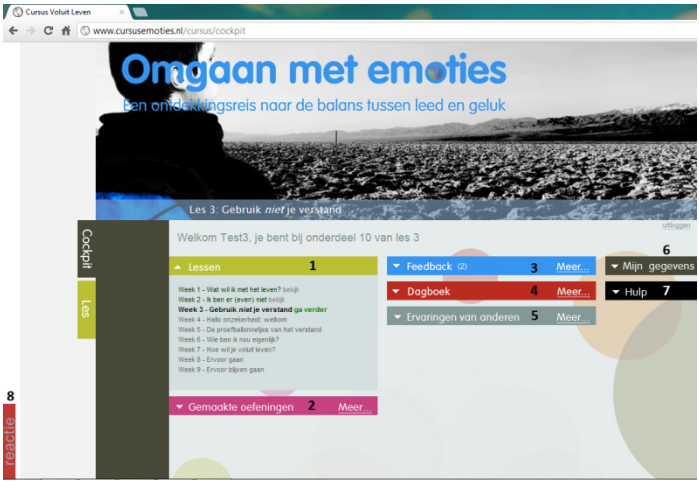


Fig. 1. Personal home screen of the web-based intervention used as an example with the elements: lessons (1), overview of completed exercises (2), feedback (3), diary (4), success stories (5), my account (6), help (7) and a 'react' button where respondents could comment on the application (8)

2 Methods

For this paper, data about usage of the web-based intervention, that has been investigated in a different study, has been taken as an example. In-depth information about the study and the methods can be found in Kelders [12].

This usage data was automatically collected by the intervention itself in the form of log-files. The log-files contained a record of actions taken by each participant with for each action the following information: user-id; action type (for example 'logged in') ; action specification (for example '1-11' [lesson 1, screen 11]); time and day. An excerpt of this usage data can be found in table 1.

Table 1. Excerpt of example log-files, with on each row from left to right: user-id; action type; action specification; time and day.

user-id	action type	specification	time and day
1096	logged in	1-11	30-3-2011 11:35
1096	viewed text-message	2	30-3-2011 11:35
1096	logged out	1-11	30-3-2011 11:35
1081	logged in	0-0	30-3-2011 11:39
1064	logged in	2-1	30-3-2011 11:49
1064	viewed text-message	3	30-3-2011 11:49
1064	logged out	2-1	30-3-2011 11:50
986	logged in	2-1	30-3-2011 11:59
986	logged in	2-1	30-3-2011 12:21
1093	logged in	2-1	30-3-2011 12:25
798	logged in	1-12	30-3-2011 12:27
798	viewed text-message	1	30-3-2011 12:27
798	logged in	1-12	30-3-2011 12:28
1072	logged in	1-1	30-3-2011 12:28
798	logged out	1-12	30-3-2011 12:31
898	logged in	2-1	30-3-2011 12:58
812	logged in	1-1	30-3-2011 12:59
898	logged in	2-1	30-3-2011 13:01
1079	logged in	0-0	30-3-2011 13:09
812	logged out	1-1	30-3-2011 13:10
895	logged in	2-1	30-3-2011 13:12
814	viewed success story	15	24-4-2011 21:49
818	logged in	2-1	30-3-2011 13:13

From this 'raw' usage data, the number of times each participant performed an action in the web-based application was extracted. Using the 'action specification' data, these actions were further classified. Following are examples of the actions 'login' and 'viewing success stories'. For logins this meant that not only the total number of logins per participant during the intervention period was extracted, but furthermore, the number of logins per participant per lesson of the intervention. Moreover, from this data it was extracted which lesson each participant had reached. For the viewing success stories, not only the number of times a success story was clicked was extracted, but also how many and what percentage of possible unique stories were clicked.

For a more in-depth pattern analysis, we arbitrarily selected 20 participants (5 participants that reached lesson 3 or 4; 5 participants that reached lesson 6 or 7; and 10 participants that completed the full intervention). Effort was made to ensure that selected participants had the same distribution of demographic characteristics and randomized group as the full sample. Of these participants, we examined all actions in lesson 2 (all selected participants), lesson 5 and lesson 8 (for those that reached these lessons) to identify emerging use patterns. We chose to examine these lessons because they reflect the three main segments of the content of the intervention and because we wanted to avoid the first and the last lesson for the expected non regular use pattern in these lessons. We expect the participants to explore and get to know the application more in the first lesson and the last lesson is shorter (i.e. less text and exercises) than the other lessons. For each participant and for each of the selected lessons, we recorded all actions in between the time they started the lesson under investigation and the time they started the following lesson.

We chose to do this analysis only for a small subsample of the data, because the focus of this analysis was on pattern recognition related to use of features of the interventions. Furthermore, the choice was pragmatic, because due to the lack of software to analyze logfiles at that time, all analyses were done by hand.

3 Results

The full results of the example study are presented elsewhere [12]. However, some of the data will be presented here as an example of the results that can be attained using log-data.

First, from the log-data on the lesson reached per participant, an adherence or attrition diagram can be attained. Figure 2 presents the attrition diagram for the 206 participants that started lesson 1 in the aforementioned study of the web-based intervention "Living to the Full" [12].

The average number of lessons started was 6.9 out of a possible 9 and 57% of the participants in this study completed all lessons. From this diagram we can see that the largest group of non-adherers started to non-adhere during lesson 2, followed by lessons 3 and 6. This is important to know at *what* moment persuasive features can be of value to motivate persistence.

Second, the extraction of the number of times each participant performed an action in the web-based application yields, for example, the results shown in Table 2.

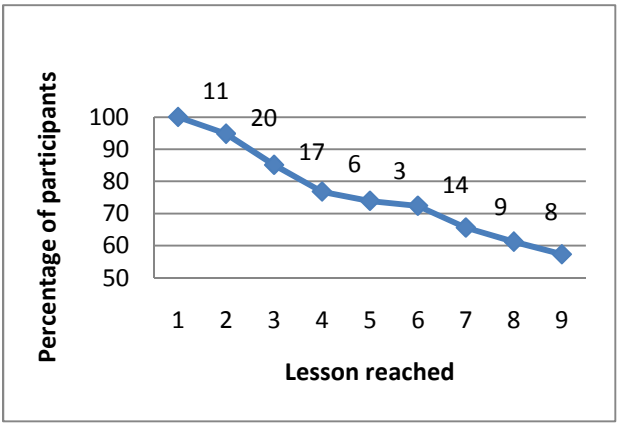


Fig. 2. Example attrition diagram. Highest lesson reached by percentage of participants, with above each line the number of participants that started to non-adhered during that lesson [12].

Table 2. Example actions per participant during the whole intervention period

Action	Average number of times per participant
Login[a]	
total, n	29.1
per started lesson, n	3.9
Feedback messages viewed	
total, n	15.7
unique messages, n	8.5
unique messages per lesson, n	1.1
Mindfulness exercises	
total started, n	6.0
unique started, n (%)[b]	2.9 (73.0)
unique downloaded, n (%)[b]	1.9 (45.6)
unique used, n (%)[b]	3.5 (85.0)
Success stories viewed	
total, n	6.5
unique, n (%)[b]	4.0 (59.1)

[a] Only logins more than 30 minutes after the previous login were counted
[b] % = unique actions / possible actions

Third, the pattern analysis yielded, for example, the use patterns of 5 respondents for lesson 2 (Table 3). From these patterns, critical episodes for persuasive features to support usage of content can be identified. Examples of these critical episodes are:

- Sessions that involve only a login and a logout action, with less than a minute in between
- Feedback messages are not read the first session after they are available

Table 3. Example use patterns of 5 respondents for lesson 2

Session	User 1	User 2	User 3	User 4	User 5
1	login; start lesson	login; start lesson	login; start lesson, mf player 2; multimedia, logout	start lesson; download mf 2	start lesson; login; mf player 2; logout
2	login; continue lesson; multimedia; mf player 2; download mf 2; feedback 1877	login (3x); success story 2, login	login; feedback 3371	login; continue lesson; mf player 2 (2x); logout	
3		login; continue lesson		login; feedback 1; success story 2	
4		login; feedback 2106		login	
5				login; feedback 2	

4 Discussion

4.1 Implications for Persuasive Design

The results show several critical episodes for the employment of persuasive components to increase adherence during an intervention period. In depth analyses of the results showed when and for whom persuasive components can matter.

The logdata revealed that not all content features (like feedback, success stories) were used. The overall study showed that users (adherers and non-adherers) differ in the usage of the content (program, lessons to be completed) [12]. To increase the adaptability of the intervention to users and usage, continuous process measurement should be built in the system and scheduled during treatment process instead of a fixed measurement approach (before and after the usage of an intervention).

The results indicate that there are critical moments for persuasion. Critical episodes for employment of persuasive features in this mental health therapy setting are the focus on the willingness to follow a therapy, the awareness of their non-coping strategies, and the adoption of "new" skills for behavior change. From prior research [27] we know that in other interventions (lifestyle program) 3 weeks after start might be a critical point for drop-out. Saturation (attrition curve [28]) can be a reason, as well as characteristics of the therapeutic program. In any case, it shows that technology can

be supportive to design persuasive content. Persuasive features are needed at those critical episodes to support motivation, to explain the risks of drop out, and to stimulate users in a positive way to achieve their goals. The differences in user patterns can be used to motivate low adherers to login and to remind them to use the content that is crucial for progress in treatment. The effectiveness of reminders can be evaluated to log the moments of prompting and the actions taken based upon those reminders.

Remarkable is that, in this show-case, feedback was not checked when this is available. One reason could be that participants are not triggered to use feedback or that feedback is not perceived or experienced as important for progress and persistence. The role of feedback, human or automated, should be further investigated. We know from the other experiments with the same intervention that although human support is more effective during the intervention, automated feedback is as effective as human support three months after the intervention has been completed [12]. The effects of mode and frequencies of interactions (feedback, human, automated) with a system should be studied more in depth to know how dialogue support can be optimized. Especially automated feedback can be of value for reasons of efficiency.

A different possibility to adapt the design of web-based interventions to enhance adherence can be found in use patterns. It has been observed that adherers and non-adherers seem to have different use patterns. With this insight, it is possible to act: either on the level of the intervention or on the level of the participant. On the level of the intervention it may sometimes be advantageous to adapt to the observed patterns (i.e. 'pave the cowpaths': "look where the paths are already being formed by behavior and then formalize them, rather than creating some kind of idealized path structure that ignores history and tradition and human nature and geometry and ergonomics and common sense" [29]), while on the other hand, it may sometimes be more prudent to adapt the design of the intervention to increase the likelihood of preferred patterns. On the level of the participant, the intervention can be designed to provide guidance to the participant to increase the likelihood of employing a preferred pattern (e.g. by providing feedback on the usage pattern or providing links to features that should be accessed on the home page) or intervene when the chance of the participant becoming a non-adherer is high, either through the intervention itself or through, for example, a phone call from a care professional. Adaptive interfaces seem to provide a method of achieving this flexibility. When redesigning an intervention in this manner, attention should be paid not only to the system, but also to the content as it is likely that changing one aspect influences the other aspects. Losing the 'holistic' overview may lead to a deterioration of usage and adherence instead of the desired improvement [2].

5 Conclusion

The capacities of technology to persuade should be articulated in the development of web-based interventions to change behaviors. Built in log data and user data (demographics; eHealth literacy, involvement etc.) can provide information to know when and for whom persuasive features are important. Further research is needed on what features, dose and timing of persuasive features matter most for adherence and for the effects to be achieved. An important step might be to identify use patterns that are

related to adherence and to (re)design interventions in such a way to promote these use patterns. A different area of future research lies in the investigation of a more pragmatic way to identify participant characteristics that may influence or predict adherence, following the 'persuasion profiling' approach [30]. Furthermore, our results indicate that the content of lessons may need a different amount or mode of interaction. Here lies an interesting line of research; how can technology be supportive to present the content of eHealth interventions in a persuasive format? In ongoing research projects we are investigating how different persuasive formats can increase adherence and effects of web-based systems aimed at behavior change, using the persuasive system design model [31].

Analyzing logfiles is time consuming; to use the results of real-time usage during an intervention period to prompt persuasive triggers in time and in the right format, we need more advanced analytics. More advanced analytics are also needed to identify patterns that emerge and that can further clarify the in depth interplay between content an system.

References

1. Nijland, N., van Gemert-Pijnen, J.E.W.C., Kelders, S.M., Brandenburg, B.J., Seydel, E.R.: Factors Influencing the Use of a Web-Based Application for Supporting the Self-Care of Patients with Type 2 Diabetes: A Longitudinal Study. Journal of Medical Internet Research 13(3) (July-September 2011)
2. Van Gemert-Pijnen, J.E., Nijland, N., Van Limburg, M.A.H., Ossebaard, H.C., Kelders, S.M., Eysenbach, G., Seydel, E.R.: A holistic framework to improve the uptake and impact of eHealth technologies. Journal of Medical Internet Research 13(4) (2011)
3. Black, A.D., Car, J., Pagliari, C., Anandan, C., Cresswell, K., Bokun, T., McKinstry, B., Procter, R., Majeed, A., Sheikh, A.: The impact of eHealth on the quality and safety of health care: a systematic overview. PLoS Med. 8(1), e1000387 (2011)
4. Christensen, H., Reynolds, J., Griffiths, K.M.: The use of e-health applications for anxiety and depression in young people: challenges and solutions. Early Intervention in Psychiatry 5(suppl. 1), 58–62 (2011)
5. Riper, H., Smit, F., Van der Zanden, R., Conijn, B., Kramer, J., Mutsaers, K.: E-mental health. High tech, high touch, high trust. Programmeringsstudie e-mental health in opdracht van het ministerie van VWS. Trimbos-Instituut, Utrecht, The Netherlands (2007)
6. Glasgow, R.E.: eHealth evaluation and dissemination research. Am. J. Prev. Med. 32(5), S119–S126 (2007)
7. Hesse, B.W., Shneiderman, B.: eHealth research from the user's perspective. Am. J. Prev. Med. 32(5), S97–S103 (2007)
8. Pagliari, C.: Design and evaluation in eHealth: challenges and implications for an interdisciplinary field. J. Med. Internet Res. 9(2), e15 (2007)
9. Christensen, H., Griffiths, K.M., Farrer, L.: Adherence in internet interventions for anxiety and depression. Journal of medical Internet Research 11(2), e13 (2009)
10. Brouwer, W., Kroeze, W., Crutzen, R., de Nooijer, J., de Vries, N.K., Brug, J., Oenema, A.: Which intervention characteristics are related to more exposure to internet-delivered healthy lifestyle promotion interventions? A systematic review. Journal of Medical Internet Research 13(1), e2 (2011)

11. Schubart, J.R., Stuckey, H.L., Ganeshamoorthy, A., Sciamanna, C.N.: Chronic health conditions and internet behavioral interventions: a review of factors to enhance user engagement. Comput. Inform. Nurs. 29(suppl. 2), TC9–TC20 (2011)
12. Kelders, S.M.: Understanding adherence to web-based interventions. Dissertation University of Twente (2012)
13. Balmford, J., Borland, R., Benda, P.: Patterns of use of an automated interactive personalized coaching program for smoking cessation. J. Med. Internet Res. 10(5), e54 (2008)
14. Binks, M., van Mierlo, T.: Utilization patterns and user characteristics of an ad libitum Internet weight loss program. Journal of Medical Internet Research 12(1), e9 (2010)
15. Couper, M.P., Alexander, G.L., Zhang, N.H., Little, R.J.A., Maddy, N., Nowak, M.A., McClure, J.B., Calvi, J.J., Rolnick, S.J., Stopponi, M.A., Johnson, C.C.: Engagement and Retention: Measuring Breadth and Depth of Participant Use of an Online Intervention. Journal of Medical Internet Research 12(4), 41–55 (2010)
16. Davies, C., Corry, K., Van Itallie, A., Vandelanotte, C., Caperchione, C., Mummery, W.K.: Prospective associations between intervention components and website engagement in a publicly available physical activity website: the case of 10,000 Steps Australia. Journal of Medical Internet Research 14(1), e4 (2012)
17. Linke, S., Murray, E., Butler, C., Wallace, P.: Internet-based interactive health intervention for the promotion of sensible drinking: patterns of use and potential impact on members of the general public. Journal of Medical Internet Research 9(2), e10 (2007)
18. Zbikowski, S.M., Hapgood, J., Smucker Barnwell, S., McAfee, T.: Phone and web-based tobacco cessation treatment: real-world utilization patterns and outcomes for 11,000 tobacco users. Journal of Medical Internet Research 10(5), e41 (2008)
19. Krukowski, R.A., Harvey-Berino, J., Ashikaga, T., Thomas, C.S., Micco, N.: Internet-Based Weight Control: The Relationship Between Web Features and Weight Loss. Telemed. J. E-Health 14(8), 775–782 (2008)
20. Hayes, S.C., Strosahl, K., Wilson, K.G.: Acceptance and Commitment Therapy: An experiential approach to behavior change. Guilford Press, New York (1999)
21. Kabat-Zinn, J.: Full catastrophe living: using the wisdom of your body and mind to face stress, pain, and illness. Delacorte Press, New York (1990)
22. Kabat-Zinn, J.: Wherever you go, there you are: Mindfulness meditation in everyday life. Hyperion, New York (1994)
23. Bohlmeijer, E.T., Hulsbergen, M.: Voluit Leven. Mindfulness of de kunst van het ervaren, nu als praktisch hulpboek [Living to the full. Mindfulness or the art of acceptance, now as a practical help book]. Boom, Amsterdam (2008)
24. Bohlmeijer, E.T., Fledderus, M., Rokx, T.A., Pieterse, M.E.: Efficacy of an early intervention based on acceptance and commitment therapy for adults with depressive symptomatology: Evaluation in a randomized controlled trial. Behav. Res. Ther. 49(1), 62–67 (2011)
25. Fledderus, M., Bohlmeijer, E.T., Pieterse, M.E., Schreurs, K.M.: Acceptance and commitment therapy as guided self-help for psychological distress and positive mental health: a randomized controlled trial. Psychol. Med., 1–11 (2011)
26. Fledderus, M., Bohlmeijer, E.T., Smit, F., Westerhof, G.J.: Mental health promotion as a new goal in public mental health care: a randomized controlled trial of an intervention enhancing psychological flexibility. Am. J. Public Health 100(12), 2372 (2010)
27. Postel, M.G., de Haan, H.A., ter Huurne, E.D., van der Palen, J., Becker, E.S., de Jong, C.A.J.: Attrition in Web-Based Treatment for Problem Drinkers. J. Med. Internet Res. 13(4), e117 (2011)
28. Eysenbach, G.: The Law of Attrition. J. Med. Internet Res. 7(1), e11 (2005)

29. Crumlish, C., Malone, E.: Designing social interfaces, 1st edn. O'Reilly Media, Beijing (2009)
30. 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)
31. Oinas-Kukkonen, H., Harjumaa, M.: Persuasive Systems Design: Key Issues, Process Model, and System Features. Communications of the Association for Information Systems 24, 485 (2009)

Where to Place My Ambient Persuasive Display? Insights from a Six-Month Study

Patricia M. Kluckner, Astrid Weiss, and Manfred Tscheligi

HCI & Usability Unit, ICT&S Center, University of Salzburg
{patricia.kluckner,astrid.weiss,manfred.tscheligi}@sbg.ac.at

Abstract. How does the placement of an ambient persuasive display for modifying energy consumption in the household change over a period of six months? There is limited knowledge about where individuals place such displays in their private households. Location is essential for the sustainable usage of persuasive displays. There is a need to gain insights into the placement decisions of users. We gathered this information in two studies. First, we did a requirement study to collect data where participants would potentially place a display. Second, we conducted a long-term household study to review the actual placement. Participants sent us pictures of their positioning at the beginning and the end of the study. It revealed minimal changes in the position of the displays, but that the choice of position is a very intentional one. We point out our findings and the benefit of this knowledge for the development of ambient persuasive displays.

1 Introduction and Motivation

A lot of research in the domestic context exists in which the placement of information technology is explored [6, 7, 8, 10, 11]. The aim of theses studies is to identify opportunities for new technologies to be an integral part of everyday life of people in their homes. We know, from ethnographic studies in the home, that people have already become used to integrate novel information technologies in their routines and habits. However, the management of digital information and integrating it into the home is still challenging. Crabtree et al. [6] show that communication media and artifacts move from one place to another in the home, when people interact with it. We can differ between (1) ecological habitats, where artifacts live; (2) coordinate displays, where artifacts are left for others; and (3) activity centers, where artifacts are worked with.

Several studies have already explored how information technology is distributed throughout the entire home. These studies are mainly interested to find out how the placement relates to the social organization and activities of a household [7, 9, 10], but not exclusively. Tolmie et al. [5] used a video study to gather information about the placement of a new bought technology at home. According to this study, type and usage of the technology inform the decision for the arrangement in the home, which is made once in the home after the purchase. However, this decision is not always

S. Berkovsky and J. Freyne (Eds.): PERSUASIVE 2013, LNCS 7822, pp. 110–115, 2013.

stable, as e.g. other family members, who were not involved in the purchase, would consider other locations. Sometimes the positioning of the technology is a rational decision instead of a practical one, which consists of a set of components. The placement is essential for understanding the real world. While placing the technology, a social situation is created.

But is this also true for ambient displays, which should be seamlessly integrated into the home? Hazlewood et al. [12] wrote "ambient information systems describe a large set of applications that publish information in a highly nonintrusive manner, following on from Mark Weiser's concept of calm technology". This fact leads to the question of how to best embed information into the surrounding. One of their foci was where should ambient systems be placed to improve their chances of being used, without becoming distracting or annoying. Eliot et al. [11] propose flexible ambient displays, which should serve as flexible information source. This means that the information displayed is mapped to their placement. However, the goal of an ambient persuasive display is to present tailored information, following different persuasive strategies for changing people's behavior. Therefore, we argue that a constantly changing placement of an ambient persuasive display may limit its effect of sustainable behavior change.

Our own research interest lies in answering the question where people place an ambient persuasive display, which should persuade them to change their energy consumption behavior – the FORE-Watch (Forecast Of Renewable Energy-Watch, see next section for details). The research question is how does the placement of an ambient persuasive display for modifying energy consumption in the household change over a period of six months? We assumed that for an ambient persuasive display, it is relevant that the placing does not change too often in order to not lower its persuasive effect.

In the following, we present the FORE-Watch in detail, followed by a description of a co-design workshop to inform the design with respect to placement of the FORE-Watch. Next we present a six-month in-situ study with 24 households, in which we explored where participants actually positioned the FORE-Watch and if the position changed over time.

2 The FORE-Watch

The FORE-Watch, can be considered as an information compression tool; we involved forecast data and the correlation of actions and smart meter data of the feedback output. We focused on how the behavior change can be triggered through persuasive strategies via the FORE-Watch in their households.

The display works with visual cues and monitors information about the current and future state of load profile/ green power (forecast). Additionally, it provides feedback on the energy consumption as an educational function (see Fig. 1). In Figure 1, the forecast view (pictures 1 and 3) and the feedback view (pictures 2 and 4) of the FORE-Watch is depicted.

(1) (2) (3) (4)

Fig. 1. Grid status forecast (1) with the corresponding feedback (2): Tuesday, 09.04.2012 at 5:45 p.m.; and green energy forecast (3) with the corresponding feedback (4) at Thursday, 08.30.2012; 5:37 p.m.

More on the design of the FORE-Watch can be found in [3]. In the following we will present results from a co-design workshop with which we gathered data on how the display should be designed and where participants would place it in their homes (more details on the workshop can be found in [3]).

3 Co-design Workshop

The co-design workshop was based on the inspiration card technique [2], which is very well established in sustainability research [1]. One question in the workshop was where ambient persuasive information on energy saving should be placed in the home. We gave each of our six participants (p) six so-called area cards for the 'perfect' position; three cards had pictures of places in the kitchen, the foyer, and the living room and the other three cards were blank for own ideas. Additionally, we offered cards with five already developed devices for awareness of energy consumption, the Orb [13], the Flowerpot [14], the Wilting Flower [14], the Energy Aware Cord [15], and the Wattson [16].

As a result of the workshop it turned out the visualization device "Energy Aware Cord" was preferred for the living room. The Wattson [16] was considered valuable for the foyer since it is visually inconspicuous while still remains in memory to help save in a conscious way. The Flowerpot [14], the Wilting Flower [14], and the Orb [13] were considered as not advisable, because they are too conspicuous and it is not possible to put them everywhere in the home area. They are convenient, but visually not attractive. The location for the display differs from the location of visualization devices. The 'best' place for a new ambient persuasive display to present energy consumption is the kitchen or the foyer. The kitchen was mentioned as "the common room and social area where people come together, spend a lot of time and meet each other more than twice a day (p2)".

The preferred place within the kitchen was the front-side of the refrigerator. Another preferred place for the device is in the foyer as "the core of the matter and the place where household members enter or exit the home area (p5)". The usage of an

application as an information tool, which could be integrated into a clock, was a further idea from two of the three groups. In summary, the visualization devices were preferred in the living room or in the foyer and a display in the kitchen. That supports the findings of [5] and [6] that the placement of information depends on the purpose and especially, on the basis of the supported element, as a tablet or other devices, where information is anchored.

4 In-situ Six-Month Study

In a subsequent six-month in-situ study with 24 participating households we wanted to investigate where participants actually place the FORE-Watch and if the positioning changes over time. Therefore, we asked participants at the first (Trial 1) and last month of the study (Trial 2) to send a picture where they placed the watch and we added a five-item online-questionnaire to substantiate our data for both trials.

For the **first trial**, we had a response rate of almost 76%, (n=19; N=25). The kitchen was preferred by 12 inhabitants, 6 participants chose the living room and 1 participant mentioned the foyer of his flat (see Fig. 2). In all placements (kitchen, living room, and foyer), the display was placed close to another information tool or gadget and stood on a shelf (e.g., a book shelf, a commode, a table or a worktop). In the kitchen, the device was placed next to the thermostat display (2), next to the radio (3), next to a calendar (1), the television (1), the microwave (1), and the stove (1). In the living room, the situation was different; four images displayed the device close to the TV and one picture showed it between the TV and telephone. The display was also located next to the radio or to the telephone in two different occasions. One participant put his display in the middle on the dining table and one person on the kitchen countertop. In summary, the first trial of the requirement analysis identified the ideal placement for the tablet was the kitchen, followed by the living room.

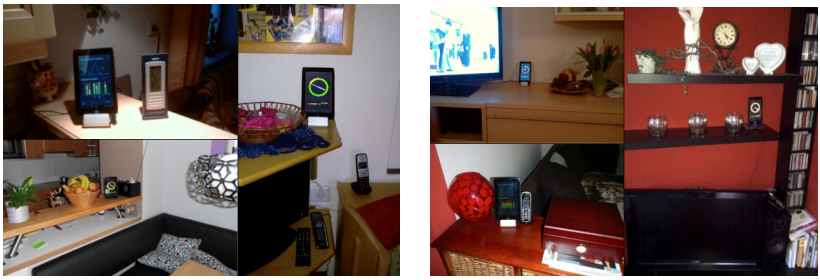


Fig. 2. Location of the display: in the kitchen area (left) and in the living room area (right)

In the **second trial**, we wanted to know if the inhabitants changed the position of the display six months later. Therefore, we asked for another picture and achieved a response rate of 67% (n=16; N=24). After analyzing the images, we had the following results. In total 13 participants put the display of the FORE-Watch Application in the kitchen and 3 in the living room. In summary, these results showed that the kitchen as the preferred area of the ambient persuasive display did not change by the end of the study. In order to gain more details about their process of decision-making for the task

of positioning the tablet, we created a short questionnaire with five open-ended questions (Q1-5). The results can be summarized as follows. (Q1) On average, 2 people decided the place of the FORE-Watch. The involvement of the children was very low, 4 out of 27 children were present during the decision process. (Q2) Only one place for location was discussed in 13 households, 5 households decided between two positions and 1 household discussed more options. Regarding Q3 (Where did you place the display in the beginning of the study?), was answered by 13 participants with kitchen, 5 in the living room, and 1 in the foyer. These statements did not change after six months, where we asked of the current place of the tablet (Q4). During the six months, 1 participant changed the position from the kitchen to the living room, 16 changed nothing, 1 put it half a meter on to the right for a better perception, and 1 participant moved it from the position next to the TV to a shelf above the TV (Q5).

In general, our analysis of the five-item questionnaire indicated that no difference could be identified in the answers compared to the actual positioning of the ambient persuasive display at the beginning and at the end of the study. Furthermore, if we look at the data of "sent pictures", we also found no differences between these measuring times. When we analyzed the change of the placement over time in the five-item questionnaire (Q5), we saw that there was a marginal difference over time; a deviation of almost 16% exist, as 3 out of 19 participants changed the location of the FORE-Watch. The preferred area in the household was more than 68% for the kitchen (69% in Trial 1; 75% in Trial 2). This result confirmed our design decision derived in from the co-design workshop, to develop the ambient persuasive display as a digital "kitchen-clock". The participants obviously understood an accepted this design decision, as the kitchen was predominantly chosen as place, without us giving any indication for this. The most popular positions were the windowsill and the place where other infotainment devices were arranged.

Due to these findings we can answer our research question on how the placement of the ambient persuasive display FORE-Watch changes with: Rarely. According to our results, only a little percentage of participants changed the position in six months and only 1 participant really put it into another room.

5 Discussion and Conclusion

We began our research with the question of where people would place an ambient persuasive display for changing energy consumption in the household and if the placement may change over time. The findings of our household study showed that the informed design of the FORE-Watch as a kitchen clock worked out to communicate the usage intent for the ambient persuasive display. Only 16% of participants changed the position of the display during the in-situ study. Our results thereby go in line with [5] and [6] that the decision where to place an information display is made once in the beginning and rarely changed, but that it can be a process in some households to optimize positioning. These results a favorable for ambient persuasive displays, as we argue that a constantly changing placement of an ambient persuasive display may limit its effect of sustainable behavior change. However, we have to state one limitation, namely that that the tablet PC we used for the ambient display needed a constantly electrical socket for loading the device.

In conclusion, our exploratory research gave interesting insights in the location-finding process for the sustainable usage of an ambient persuasive display in the home. We could show how essential it is to involve future users not only in the display design process, but also in the affordance design in terms of placement. Moreover, we demonstrated the importance of long-term studies to understand placement requirements of ambient persuasive displays, as expectations and actual usage does not necessarily need to be the same.

Acknowledgement. The financial support of the "PEEM Project" by the Austrian Research Promotion Agency (FFG) is gratefully acknowledged. We also thank our project partners Salzburg AG and the Center for Usability and Engineering (CURE).

References

1. Davis, J.: Participatory design for sustainable campus living. In: Proc. of CHI EA, pp. 3877–3882 (2010)
2. Halskov, K., Dalsgard, P.: Inspiration card workshops. In: Proc. of DIS, pp. 2–11 (2006)
3. Schrammel, J., Gerdenitsch, C., Weiss, A., Kluckner, P.M., Tscheligi, M.: FORE-Watch – The Clock That Tells You When to Use: Persuading Users to Align Their Energy Consumption with Green Power Availability. In: Keyson, D.V., Maher, M.L., Streitz, N., Cheok, A., Augusto, J.C., Wichert, R., Englebienne, G., Aghajan, H., Kröse, B.J.A. (eds.) AmI 2011. LNCS, vol. 7040, pp. 157–166. Springer, Heidelberg (2011)
4. Staats, H., Harland, P., Wilke, H.A.M.: Effecting Durable Change. Environment and Behavior 36(3), 341–367 (2004)
5. Tolmie, P., Crabtree, A.: Deploying Research Technology in the Home. In: CSCW, pp. 639–648 (2008)
6. Crabtree, A., Rodden, T.: Domestic Routines and Design for the Home. In: CSCW, pp. 191–220 (2004)
7. Elliot, K., Neustaedter, C., Greenberg, S.: Time, Ownership and Awareness: The Value of Contextual Locations in the Home. In: Beigl, M., Intille, S.S., Rekimoto, J., Tokuda, H. (eds.) UbiComp 2005. LNCS, vol. 3660, pp. 251–268. Springer, Heidelberg (2005)
8. Hindus, D.: The Importance of Homes in Technology Research. In: Streitz, N.A., Siegel, J., Hartkopf, V., Konomi, S. (eds.) CoBuild 1999. LNCS, vol. 1670, pp. 199–207. Springer, Heidelberg (1999)
9. Taylor, A., Swan, L.: Artful Systems in the Home. In: Proc. of CHI 2005, pp. 641–650. ACM Press (2005)
10. Harper, R., Evergeti, V., Hamill, L., Strain, J.: Paper-mail in the Home of the 21st Century. In: Proc. of Digital Technology in Home Environments (2001)
11. Elliot, K., Watson, M., Neustaedter, C., Greenberg, S.: Location-dependent information appliances for the home. In: Proc. of Graphics Display, pp. 151–158 (2007)
12. Hazlewood, W.R., Coyle, L., Pousman, Z., Lim, Y.-K.: Ambient Information Systems. In: Second International Workshop on the Design and Evaluation of Ambient Information System (2008)
13. Energy Orb, `http://www.ambientdevices.com/cat/orb/PGE.html`
14. Wilting Flower, Flower Pot, `http://www.treehugger.com/corporate-responsibility/the-wilting-flower-and-the-flowerpod-two-budding-concepts-for-visualizing-home-energy-use.html`
15. Gustafsson, A., Gyllenswärd, M.: The Power-Aware Cord: Energy Awareness through Ambient Information Display. In: Proc. of CHI (2005)
16. Wattson, `http://www.diykyoto.com/uk`

Lingering Serious Experience as Trigger to Raise Awareness, Encourage Reflection and Change Behavior

Tim Marsh[1] and Brigid Costello[2]

[1] James Cook University, QLD, Australia
tim.marsh@jcu.edu.au
[2] University of New South Wales, Australia
bm.costello@unsw.edu.au

Abstract. While work in interaction design, human-computer interaction (HCI) and the games literature begins to address experience beyond positive, it just scratches the surface. By turning to drama, literature, music, art and film that has shaped experiences and emotion beyond the positive and fun for many years, we describe what experience beyond positive looks like, show how it is not always "uncomfortable" and argue for the more appropriate term "serious experience". We discuss the importance of the take-away message / serious experience in persuasive technology, persuasive games and serious games to linger or resonate post-encounter for user/players to encourage reflection, affect attitudes and change behaviors in order to fulfill a persuasive purpose. Finally, we describe associated ethical concerns and make recommendations for designers, evaluators and practitioners in order to safeguard players/users.

1 Introduction

Like drama, literature, music, art and film, we argue that experience from interaction or play with persuasive technology and games shouldn't exclusively be positive nor always have a happy resolved ending. Instead we propose that persuasive technology and games may aim to fulfill their purpose by evoking less fun positive experiences. For example, where the purpose is to provoke thought, provide a message or an experience on a particularly difficult, uncomfortable or unsettling subject or issue. In addition, we argue that experience with persuasive technology and games needs to resonate or linger with the user/player after an encounter to encourage reflection, affect attitudes and change behaviors in order to fulfill a persuasive purpose.

This paper is organized as follows. We review the literature on experience in interaction, HCI and games, and identify limitations with this work. We then turn to other media and art forms including drama, literature, music, film and interactive art to show how these limitations can be addressed. Next we propose the categorization *serious experience* to encapsulate experience beyond positive and fun, and to provide necessary and appropriate experience to linger or resonate post-encounter for users/players in order to encourage reflection and change attitudes and behaviors to fulfill purpose. Finally, we briefly talk about the associated ethical concerns and make recommendations for designers, developers and practitioners in order to safeguard players.

S. Berkovsky and J. Freyne (Eds.): PERSUASIVE 2013, LNCS 7822, pp. 116–124, 2013.

2 Experience in Play and Interaction

In computer, video and digital games, *experience* has been the main driver for design since their inception. The term *player experience* is used to frame experience that players get from playing games and specifically, player experience is widely described under the broad term *fun.* According to Salen & Zimmerman [38] "Good games are fun. Fun games are what players want". Fun is "central to the process of making good games" [14]. Strong support for these claims is provided in an Entertainment Software Association[1] survey that reported 87% of the most frequent game players cited fun as the first reason why they play video games [13].

Likewise, in the design discipline of HCI, much work has attempted to shed light on the composition and foundational elements of experience and user experience in interaction design. This is demonstrated in proposed theories, levels, threads, frameworks and design research and thinking that emphasizes the pleasure [27], hedonic [19], ludic [16], emotional [35], enchantment [49] and fun and enjoyment [5] qualities and value in interaction experience.

However, HCI has struggled to reach a common understanding and consensus definition [e.g. 20, 29] demonstrating the multifarious and elusive nature of experience. While there is little doubt that work on experience in HCI has been instrumental in providing a language and in refocusing interaction and product design towards a broader experiential perspective, two major criticisms can be attributed to much of this work.

First, it is invariably restricted to positive, fun and aesthetic experience [21, 40]. Second, HCI has largely been concerned with the moment of experience and tends to ignore things that "outlive the moment experience" that people really "value" and "find worthwhile" [10].

2.1 Experience beyond Positive and Fun

We argue that for designers to focus on fun means that they might take a shallow or cursory approach to the design of negative affect rather than advocating design and development of alternative and deep experiences and emotion beyond fun in games. Consider for example a persuasive game like Darfur is Dying which has been very successful in raising awareness on the plight of starving people in Darfur and the everyday risks and fear they encounter from armed militia, and in persuading people to take action. However, as the game itself is fun to play, it can create a level of unease because of the tension between, on the one hand, the fun gameplay, and on the other, the serious subject matter it portrays. For the informed player this tension is awkward and discomforting and can appear disrespectful of the suffering, misery and distress endured by the people the game is intended to help.

Ian Bogost, co-founder of Persuasive Games, in an interview with Gamasutra [15] similarly argues for the development of human experiences beyond fun in games:

[1] ESA: Entertainment Software Association, US trade association for video games whose members include Atari, Electronic Arts, Microsoft, Square Enix et al.

> *"For 30 years now we've focused on making games produce fun" "Isn't it about time we started working toward other kinds of emotional responses?" "I know that comparisons to the film industry have grown tired and overused," he says, "but indulge me in this one: When you watch the Academy Awards this year, how many films in the running for awards are about big explosions and other forms of immediate gratification, and how many are about the more complex subtleties of human experience? "Someday, hopefully someday soon, we'll look back at video games and laugh at how unsophisticated we are today".*

The games literature is increasingly identifying that designing exclusively for the experience of fun in games is too limiting. Hunicke, LeBlanc and Zubek [24] argue for a move away from words like fun towards a more appropriate vocabulary to describe "the desirable emotional responses evoked in the player, when she interacts with the game system". Similarly, Calleja [9] identifies limitations with the term fun applied to games arguing that "pinning motivation for game-playing on the notion of fun risks missing important dimensions of the game experience". While work in games proposes experience beyond the positive, for example as shown in the claimed "comprehensive categorization of digital game experience" [36], with two categories, out of nine being associated with "negative experience" (*negative* affect: frustration, disappointment, irritation, anger; *suspense*: challenge, tension, pressure, hope, anxiety, thrill), the title however, appears to identify the authors' point of view that whether or not game experience is positive or negative, "[i]t is always a lot of fun!".

So while the HCI community of ACM SIGCHI finally opens-up shop on experience other than positive, fun and aesthetic, as elegantly captured in Benford et al.'s [4] work on "uncomfortable interaction", and the game literature increasingly identifies the importance of moving beyond the fun game experience, much work is still largely tentative, and is only just beginning to scratch the surface.

Looking to other media, drama, music, art and film that provide powerful and deep experiences and emotion to inform user and player experience provides some leverage to these discussions. In music compositions that create variations in feelings, moods and emotions beyond the positive – it would be quite limiting and tedious if all music was restricted to just positive and fun. In drama, literature, film and storytelling in general, experience beyond positive is necessary to portray suffering, struggle, conflict and adversity, etc. For example, in typical drama and story structure such as the 3 or 5 act play, to set-up a rising action or conflict, that is typically followed by a resolution, but not necessarily a pleasurable one (e.g. Shakespearean tragedies). Grodal [18] has looked to film experience in an attempt to understand video game experience and building on Zillmann's [43] work on the psychology of suspense in drama and film, Klimmt et al. [28] identified suspense in video games.

In addition, much work in HCI and games adopted Boorstin's [6] three Vs foundational elements of experience and emotion from film: voyeuristic (new and the wonderful), visceral (thrills, spectacle and suspense) and vicarious (empathy and emotional transfer). In HCI and interaction design, the three Vs also played a prominent role in proposals for shifts "from usability to user experience" [42] and in informing underlying foundational elements for experience and emotion in interaction

and product design in popular HCI texts [32, 35]. However, again the emphasis in this work has been on positive and fun experience to inform HCI, but which largely disregards the power of the three Vs to describe experience beyond positive [33].While Norman [35] acknowledges the importance of negative emotion in design as suggested in the sub-title of his book, "Why we Love (or Hate) Everyday Things", he offers only a cursory discussion on the negative.

While the three Vs experience continues to be adopted and applied to video games, again this work largely focuses on fun experience. For example, [14] identifies the three Vs as "corresponding" to his framework of fun in "Natural Funativity's Physical, Social, and Mental fun"; and [37, 39] adopt the three Vs to help talk about the fun and experience of playing a game. However, the beauty and power of the three Vs is in its ability to frame a broad range of experience and emotion - both "positive" and "negative" (frightened, disgusted, nauseated, tense, sad, angry, weak, tension, cowardly, serious) as shown in study results from survey and interview approaches experienced by almost all players with our test education and first-person shooter games [30, 31].

Other work in games that aim to create an experience beyond fun by capturing and expressing "difficult emotions with a games mechanic", that linger after the game has finished, comes from Brenda Brathwaite and her well-known GDC 2010 talk [7] and her "works" within the Mechanic is the Message [8] series of non-digital games. These include: The New World (2008) about the Middle Passage and slave trade, Síochán leat aka "The Irish Game" (2009) about the Cromwellian Invasion of Ireland, and Train (2009) a game about the Holocaust and the transportation of people to concentration camps [8]. Participants of these games learn about, and are complicit in, difficult subjects that either emerge during gameplay or are revealed fully after the game (e.g. Train), and the associated difficult experiences and emotions linger after the game has finished.

Although often using similar interactive technology and games, interactive art has never shied away from creating uncomfortable or unpleasant experiences. Artists might use exaggeration, shock or disorientation to create experiences of alienation. They might also create works that ask their audience to subvert or resist common uses or purposes of a technology [26].

For example, in *Pin Cushion* [41] the audience is invited to distort a representation of a human female face by pricking it with large acupuncture needles. However, they have less control than they think. As a participant touches the needles the artwork reads her or his body's electrical conductivity, resistance and charge. It is this intimate reading that impacts the lifespan and well-being of the character.

Working against the usual excitement of interactive technologies, *Perversely Interactive System* [25] uses a biofeedback device to measure tension levels in the participant, with the character in the artwork only responding when tension levels are low. As the artists describe, this was uncomfortable for the participant because it meant that "getting what one desired required controlling or denying that desire". In each case, the audience unease or discomfort is used to provoke interpretative reflection.

Finally, we identify examples in interactive art and games where the user/player takes pleasure from negative experience. Early 20th century conceptions of play included experiences of physical pain and mental suffering. The pleasure that we

might take from probing a sore tooth or experiencing the sadness of a tragic artwork being described as a form of playing with emotions that stems from a need to "satisfy our craving for intense impressions" [17].

Recent frameworks from games and interaction design researchers also include experiential categories that go beyond common conceptions of fun. For example, Bartle's [3] model of player types in MUDs includes *killers*, a type of player who derives pleasure from bullying and/or manipulates others.

Costello and Edmond's [11, 12] pleasure framework includes the category of *subversion*, which describes the pleasure that can be had by behaving against the norm, by breaking rules or of seeing others break them. Building on this framework and with a focus on game experience, Arrasvuori et al [2] have added the categories of *cruelty* and *suffering*. Cruelty is the playful experience of acting to cause physical or mental pain in others. Suffering they describe as encompassing the emotions of "boredom, stress, anxiety, anger, frustration, loss and even humiliation".

While these examples demonstrate that experiences and emotion beyond the positive from interaction and play is an area that continues to be enthusiastically explored in games and interactive art, the negative and potentially extreme experiences from encounters suggests that precautions must be taken to ensure the safety and well-being of players/users.

3 Serious Experience

As shown in table 1, we argue that experience from an encounter (interaction or play) with persuasive technology, persuasive and serious games is framed within two main categories: positive and serious; and propose that design should be an appropriate rhythm between these two.

Table 1. Between Positive and Serious Experience

<table>
<tr><th colspan="3">Experience</th></tr>
<tr><th>Positive</th><th colspan="2">Serious</th></tr>
<tr><td rowspan="2">Fun</td><td>Thought-Provoking</td><td rowspan="2">Negative, Uncomfortable, Unpleasant, Provoking</td></tr>
<tr><td>Positive-Negative</td></tr>
</table>

Serious experience encapsulates experience beyond positive and fun, and is composed of two sub-categories. The first category generally identifies experience that is neither exclusively positive nor negative/uncomfortable, but falls somewhere in-between. These are entertaining, likable, or where user/player takes pleasure from

negative experience. For example, interaction or play that is thought-provoking, informing, raises awareness on issues, or where the user/player takes pleasure from negative experience, variously described in interactive art as "pleasurable sense of unease" and "pleasurable thrill of danger" [12], in games as "positive negative experience" [23, 34], and in learning games as "pleasant level of frustration" [22]. This categorization of experience is *entertaining without being exclusively fun*.

The second category is "uncomfortable" and "negative experience", as discussed above. These extreme experiences and emotions are disturbing, discomforting and provoking, and in persuasive technology and games the user/player unease or discomfort is used to provoke interpretative reflection and encourage changes in behaviors in order to fulfill a persuasive purpose.

We acknowledge that an encounter with persuasive technology and game may be experienced differently at different times by the same user/player or can be experienced differently by different users/players. This depends not on the experience itself but on the perception of the person who experiences it. For example, fun at one time and thought-provoking the next or one person experiences a negative and unpleasant experience while another experiences it as thought-provoking. One theory that may help provide some leverage in further investigations is Apter's [1] reversal theory where the exact same type of high (or low) arousal experience could cause one person to experience it as unpleasant and the other as pleasant.

4 Ethics and Code of Practice in Serious Games

Our concern is not only with the moment-to-moment and in-game experience per se that has dominated work in video games and interaction design, but also on experience that lingers or resonates with users/players after an encounter. This is similar to the idea of bleed in games where a weakening of the protective frame of play allows emotion and experience to bleed out from the game and influence the player outside the game beyond the magic circle [28]. As it is these lingering and resonating experiences that users/players take-away that often provide a measure of success of purpose in serious games, as designers and developers we must be aware of the potential danger and harm that serious games could cause.

While drama, performance, literature and film have portrayed similar extreme and difficult topics, perhaps similar age/rating systems should be introduced. We recommend that developers of games with such extreme topics are aware of the ethics surrounding their development, that guidelines should be drawn-up to inform design and development and, in some cases, they are used only under rigorous procedures and are followed by debriefing sessions (similar to those used in psychology experiments and HCI studies) to safeguard and protect players from harm.

5 Conclusion and Future Work

As discussed, much of the literature on interaction and game experience has tended to focus on positive and fun experience. While emerging work in interaction and video

games also identifies "negative" experience as being crucial to provide deeper experience and emotions, discussions are either cursory, don't go far enough, and/or are about the temporary sensations used to set-up a rising action or conflict, and is typically followed by a more pleasurable resolution.

Like drama, literature, music, art and film, we have argued that experience from interaction or play with persuasive technology and games shouldn't exclusively be positive nor always have a happy resolved ending. Instead we are proposing that persuasive technology and games may aim to fulfill their purpose by evoking less fun positive experiences. For example, where the purpose is to provoke thought, provide a message or an experience on a particularly difficult, uncomfortable or unsettling subject or issue. In addition, we have argued that this experience needs to resonate or linger with the user/player post-encounter to encourage reflection and help change attitudes and behaviors in order to fulfill a persuasive purpose.

In order to frame experiences and emotions, we propose *serious experience* – being, thought-provoking *and* negative / uncomfortable / provoking – as well as *positive experience* (fun) are essential for informing the design repertoire for interaction and play in persuasive technology, and persuasive and serious games.

Finally, as it is important for serious experience to linger or resonate post-encounter for users/players in order to encourage reflection and fulfillment of persuasive purpose, we propose that designers, developers and practitioners are aware of the ethical concerns and content rating systems are in place in order to safeguard and protect players from harm.

We are currently exploring ways to utilize the idea of *serious experience* in simulations and games. One example is an analogue game to help participants learn about and understand the environmental conditions of the Great Barrier Reef. Typical scenarios are represented in the game mechanics and through play/interaction participants learn how sensitive ecosystems operate and the impacts humans have on them. As the game unfolds, participants become aware that their gameplay/interaction is having a harmful effect on reef and marine life. Awareness and experience of having been complicit in its destruction is intended to resonate and linger on after the game has finished. The degree to which this experience lingers and how it may change behavior to fulfill a persuasive purpose will be investigated in planned studies. On-going work is currently being undertaken to transpose the Reef awareness analogue board game into a digital persuasive gaming simulation.

References

1. Apter, M.J.: Reversal Theory: Motivation, Emotion and Personality. Routledge, London (1989)
2. Arrasvuori, J., Boberg, M., Korhonen, H.: Understanding Playfulness - An Overview of the Revised Playful Experience (PLEX) Framework. In: Proc. of Design & Emotion 2010 Conference, Design and Emotion Society (2010)
3. Bartle, R.A.: Virtual Worlds: Why People Play. In: Alexander, T. (ed.) Massively Multiplayer Game Development 2 (Game Development). Charles River Media, Inc., Rockland (2005)

4. Benford, S., Greenhalgh, C., Giannachi, G., Walker, B., Marshall, J., Rodden, T.: Uncomfortable Interactions. In: ACM SIGCHI 2012 Conference on Human Factors in Computing Systems. ACM Press (2012)
5. Blythe, M.A., Overbeeke, K., Monk, A.F., Wright, P.C.: Funology: From Usability to Enjoyment, pp. 31–42. Kluwer Academic Publishers, Dordrecht (2003)
6. Boorstin, J.: Making Movies Work: Thinking Like a Filmmaker. Silman-James Press, Beverley Hills (1990)
7. Brathwaite, B.: Train: How I Dumped Electricity and Learned to Love Design. In: GDC 2010 (2010), http://www.gdcvault.com/play/1012259/Train-(28or-How-I-Dumped
8. Brathwaite, B.: The Mechanic is the Message, http://mechanicmessage.wordpress.com/
9. Calleja, G.: Game Design Involvement A Conceptual Model. Games and Culture (2), 236–260 (2007)
10. Cockton, G.: Designing worth is worth designing. In: Proceedings of the 4th Nordic Conference on Human-Computer Interaction, NordicCHI 2006, pp. 165–174. ACM Press (2006)
11. Costello, B., Edmonds, E.: A Study in Play, Pleasure and Interaction Design. In: Designing Pleasurable Products and Interfaces, University of Art and Design Helsinki, pp. 76–91. ACM Press (2007)
12. Costello, B., Edmonds, E.: A tool for characterizing the experience of play. In: Proceedings of the Sixth Australasian Conference on Interactive Entertainment, Sydney, Australia, pp. 1–10. ACM Press (2007)
13. ESA. Entertainment Software Association, State of the Industry Report 2000-2001. Entertainment Software Association (2001)
14. Falstien, N.: Natural Funativity, Gamasutra (2004), http://www.gamasutra.com/features/20041110/falstien_pfv.html
15. Ochalla, B.: From Gamasutra Article: June 29 2007, Who Says Video Games Have to be Fun? The Rise of Serious Games (2007), http://www.gamasutra.com/view/feature/129891/who_says_video_games_have_to_be_.php
16. Gaver, B., Martin, H.: Alternatives: exploring information appliances through conceptual design proposals. In: CHI 2000: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 209–216. ACM Press, New York (2000)
17. Groos, K.: The Play of Man. William Heinemann, London (1901)
18. Grodal, T.: Video games and the pleasure of control. In: Zillmann, D., Vorderer, P. (eds.) Media Entertainment: the Psychology of its Appeal, pp. 197–214. Erlbaum, Mahwah (2000)
19. Hassenzahl, M., Platz, A., Burmester, M., Lehner, K.: Hedonic and Ergonomic Quality Aspects Determine a Software's Appeal. In: Proceedings of CHI 2000: Human Factors in Computing Systems, pp. 201–208. ACM Press, NY (2000)
20. Hassenzahl, M., Tractinsky, N.: User experience a research agenda. Behaviour and Information Technology 25(2), 91–97 (2006)
21. Hassenzahl, M., Diefenbach, S., Göritz, A.: Needs, affect, and interactive products – Facets of user experience. Interacting with Computers 22(5), 353–362 (2010)
22. Henriksen, T.D.: Dimensions in Educational Game-Design: Perspectives on Designing and Implementing Game-Based Learning Processes in the Educational Setting, Paper for Nordic Playground event (2006)
23. Hopeametsä, H.: 24 Hours in a Bomb Shelter: Player, Character and Immersion in Ground Zero. In: Montola, M., Stenros, J. (eds.) Playground Worlds, Ropecon (2008)

24. Hunicke, R., LeBlanc, M., Zubek, R.: MDA: A Formal Approach to Game Design and Game Research. In: Proceedings of the 19th AAAI Conference, Workshop on Challenges in Game AI. AAAI Press (2004)
25. Hughes, L., Laroche, S.: Perversely Interactive System. Interactive Artwork (2004), http://www.interstices.uqam.ca/en/projects/lynn-hughes/item/27-perversely-interactive-system.html
26. Jones, C.: Excerpt from "The Mediated Sensorium". In: Jones, C.A. (ed.) Sensorium: Embodied Experience, Technology, and Contemporary Art, p. 6. MIT Press (2006)
27. Jordan, P.W.: Designing Pleasurable Products: An Introduction to the New Human Factors. Taylor & Francis (2000)
28. Klimmt, C., Rizzo, A., Vorderer, P., Koch, J., Fischer, T.: Experimental evidence for suspense as determinant of video game enjoyment. Cyberpsychology 12, 29–31 (2009)
29. Law, X.E., Roto, C.V., Hassenzahl, M., Vermeeren, A.P., Kort, J.: Understanding, scoping and defining user experience: a survey approach. In: CHI 2009: Proceedings of the 27th International Conference on Human Factors in Computing Systems, pp. 719–728. ACM Press, New York (2009)
30. Marsh, T.: Presence as Experience: Framework to Assess Virtual Corpsing. In: 4th International Workshop on Presence, Temple University, Philadelphia, USA (2001)
31. Marsh, T.: Towards Invisible Style of Computer-Mediated Activity: Transparency and Continuity. Unpublished Ph.D Thesis, Human-Computer Interaction (HCI) Group, University of York, UK (2005), http://ethos.bl.uk/OrderDetails.do?uin=uk.bl.ethos.423695
32. McCarthy, J.P., Wright, P.C.: Technology as Experience. MIT Press (2004)
33. McCarthy, J.P., Wright, P.C., Wallace, J., Dearden, A.: The experience of enchantment in human-computer interaction. Personal and Ubiquitous Computing 10(6), 369–378 (2006)
34. Montola, M.: The positive negative experience in extreme role-playing. In: Proceedings of Experiencing Games: Games, Play, and Players, 1st Nordic Digra, Sweden (2010)
35. Norman, D.A.: Emotional Design: Why We Love (or Hate) Everyday Things. Basic Books, New York (2004)
36. Poels, K., de Kort, Y., Ijsselsteijn, W.: It is always a lot of fun!: exploring dimensions of digital game experience using focus group methodology. In: Proceedings of the Future Play Conference. ACM Press (2007)
37. Rusch, D.C.: Emotional Design of Computer Games and Fiction Films. In: Computer Games as a Sociocultural Phenomenon: Games Without Frontiers, Wars Without Tears, pp. 22–32. Palgrave Publishers (2008)
38. Salen, S., Zimmerman, E.: Rules of play: Game design fundamentals. MIT Press (2004)
39. Seif El-Nasr, M., Al-Saati, M., Niedenthal, S., Milam, D.: Assassin's creed: a multicultural read. Loading 2(3) (2008)
40. Van Vliet, H., Mulder, I.: Experience and Design: Trojan Horse or Holy Grail. In: Proceedings of User Experience, Second COST294 International Workshop Held in Conjunction with NordiCHI 2006, pp. 57–62 (2006)
41. Velonaki, M.: Pin Cushion, Interactive Artwork (2000), http://mvstudio.org/work/pin-cushion/
42. Wright, P.C., McCarthy, J., Marsh, T.: From Usability to User Experience. In: Computers and Fun 3, Workshop. University of York, UK (2000); Appears in: Interfaces (46), 4–5 (Spring 2001)
43. Zillman, D.: The psychology of suspense in dramatic exposition. In: Vorderer, P., Wulff, H.J., Friedrichsen, M. (eds.) Suspense: Conceptualizations, Theoretical Analyses, and Empirical Explorations. Lawrence Erlbaum Associates, Inc., Mahwah (1996)

Personalizing Triggers for Charity Actions

Judith Masthoff[1], Sitwat Langrial[2], and Kees van Deemter[1]

[1] Computing Science, University of Aberdeen, AB25 2NH, Aberdeen, UK
{j.masthoff,k.vdeemter}@abdn.ac.uk
[2] Department of Information Processing Science, P.O. Box 3000,
90014 University of Oulu, Finland
sitwat.langrial@oulu.fi

Abstract. In this paper we investigate whether there is scope for personalizing triggers in the domain of charitable action. The first of our two studies focuses on actions promoting human rights (via Amnesty International). While participants in a previous exploratory study had indicated that victim attributes (such as gender, religion, and ethnicity) would not matter at all, we found when observing participants' actions that in fact these attributes mattered greatly. Participants tended to select victim attributes similar to their own, showing a clear potential for personalization. These findings were corroborated by a further study in the area of charitable giving (using the KIVA micro-financing website). The paper also discusses implications for digital behavior intervention.

1 Introduction

People often do not act in accordance with their beliefs. For example, most people believe that exercise is good for health, yet many do not exercise regularly. Similarly, most people believe that participating in charity actions is commendable, yet few do so regularly. There may be several reasons for such disparities, such as lack of time, lack of information and forgetfulness. Our research investigates how persuasive technology may help people overcome such a disparity between beliefs and actions. Persuasive technology is technology specifically designed to change people's attitudes and/or behaviours [4, 7]. Persuasive systems aimed at changing behaviour are often called "Digital behaviour interventions", or more recently "Behavior Change Support Systems" [23]. Persuasive technology has been used for example to motivate people to quit smoking [9], walk more [14, 24], eat more healthily [6, 11], reduce consumption of fizzy drinks [20], increase sustainable transport [21], and use less energy [13].

An extensive psychological literature exists on the factors that cause people to modify their behaviour and attitudes, leading to many behaviour change theories and techniques (see surveys in [8, 12]). A potentially important mechanism identified in these theories is that of an action trigger. Fogg also mentions that a lack of well-timed triggers may result in desired behaviour not occurring [5]. A trigger can be anything which encourages or motivates an individual to perform an action. We are investigating how best to trigger people into taking charity actions and how to trigger people to gradually take part in increasingly difficult (more time consuming) actions. Amnesty

S. Berkovsky and J. Freyne (Eds.): PERSUASIVE 2013, LNCS 7822, pp. 125–136, 2013.

International was chosen as an application domain because many people support its work, making it easy to find participants with the right attitudes, and it already has actions requiring different levels of effort.

Our triggers will be reminder messages which highlight a particular Amnesty action to participate in. This paper investigates whether and how to personalize the content of these messages to the participants[1]. By personalization we mean automatically tailoring the message to the characteristics of the recipient. In particular, we are interested in whether selecting a victim with attributes similar to the participant will make it more likely that the participant will take action. In a previous exploratory study, involving a survey with 68 participants, we investigated amongst others the perceived impact of a victim's attributes (gender, religion, ethnicity, victim being a child, and profession) on the level of motivation to take part in Amnesty actions (using Likert scales). Participants claimed that victim attributes would not have much impact, with the exception of the victim being a child which 37% thought would have some impact. However, this may be due to political-correctness, or participants not being aware of what may influence them. Psychologists have long ago found that attitudinal similarity impacts interpersonal attractiveness (e.g., [1]), which suggests that people may feel more sympathy for victims of the same religion. This has become known as the similarity-attraction paradigm which suggests that similarity increases attraction [2]. Researchers interested in persuasion have also considered similarity, but found that similarity does not always serve persuasion [17]. In this paper, we will investigate how people act when they have to select a person to help, and in particular, whether they will choose to help a person similar to themselves. We will investigate the effect of simple attributes such as age, gender and religion, and a more complicated attribute namely affinity for a country.

Within computing science, there have been several investigations into personalizing interventions (and associated user modeling, see e.g. [24]). For example, the STOP system generated smoking-cessation letters personalized to the participants' smoking habits and beliefs, attempts to quit previously and general health problems [18]. The Daphne system aimed to improve people's diet by using arguments tailored to the participants' preferences and values [6]. The PORTIA system encouraged people to improve their diets by providing personalised information to develop a discrepancy between the users' goals and their existing eating habits [11]. In the real-estate domain, Carenini and Moore found that user-tailored evaluative arguments (e.g. including "the house has a good location" for a user who cares about location) increased users' likelihood to adopt a particular house compared to non-tailored arguments [3].

There has also been a substantial amount of research into personalizing messages in the health communication domain [15]. For example, Revere and Dunbar concluded that tailored interventions are effective based on their review of 37 health behaviour intervention studies [19]. Kroeze et al found a quite strong effect of tailoring when reviewing 30 studies on computer-tailored messages for physical activity and dietary behaviour change [10]. Noar et al performed a meta-analytic review of 57 studies of tailored print health behaviour change interventions, and found a positive effect of tailoring [15]. The review in [16] shows that messages have been tailored to

[1] In addition to the *content* of the message, the *source* of the message can also impact its effectiveness [22], but this it outside the scope of this paper.

many different characteristics including participants' stage of change (e.g. contemplation: acknowledging that there is a problem but not yet ready or sure of wanting to make a change), personality (e.g. self-efficacy), health condition (e.g. body mass index), and demographics (gender, age, ethnicity). However, Noar et al see the usefulness of tailoring to demographics as a way to "place information in a meaningful context" rather than to "match content to individual' [..] interests" (for which they mention tailoring to other characteristics such as stage of change) [16]. In this paper, we will consider the role of demographic characteristics on deciding the content of the message rather than just its contextualisation.

The paper is structured as follows. Section 2 reports on a study investigating whether trigger content needs to be personalized to participant characteristics. Section 3 corroborates the findings by analysing data from an existing charity website. Section 4 draws conclusions and briefly discusses future work.

2 Study 1: Is there Scope for Personalization?

In this study, we investigate whether what people claimed in our previous exploratory study was actually true, i.e., people do not care about the age, gender, religion of the victims. We suspect that people will act differently than they said. In particular, we expect that participants' gender, age, religion, and affinity with countries will influence the selection of the victim. We also wonder how participants will rate the relative importance of attributes, when they are forced to make decisions.

2.1 Methodology

Study Design. A correlational design was used: we measured the participants' attributes and the attributes they selected for the victim and investigated the correlation between these.

Correlational variables. We considered the following pairs of variables (more detail on values is given under Materials below):

- Participant's and victim's age. We used age categories rather than exact age, as the likelihood of having an Amnesty International case with a victim of exactly the same age as the participant is not that great.
- Participant's and victim's gender. We used a simple binary decision of male and female (excluding possible transgender).
- Participant's and victim's religion. We used a limited set of religions.
- Victim's country and the country with which the participant has the highest affinity. We used a limited set of countries. The participants' affinity with each country was determined by considering (1) whether this was the participant's country of origin, (2) whether they had family there, (3) whether they had friends there and (4) whether they had travelled there. We assumed that country of origin was the highest level of affinity, followed by having family there, followed by having friends there, and followed by having travelled there.

Exploratory variables. In addition, we considered the following variables for exploratory reasons:

- Victim's case category. We used a limited set of case categories.
- Participant's importance of criteria. Participants ranked the relative importance to them of the victim's attributes and case category.

Hypotheses. We formulated the following hypotheses:

- H1. There is a positive correlation between the participant's attributes and the attributes selected for the victim. More particularly:
  - H1.1. Participants' age influences the selection of the victim's age, with participants being more likely to select a victim of the same age.
  - H1.2. Participants' gender influences the selection of the victim's gender, with participants being more likely to select a victim of the same gender.
  - H1.3. Participants' religion influences the selection of the victim's religion, with participants being more likely to select a victim of the same religion.
  - H1.4. Participants' affinity with countries influences the selection of the victim's country, with participants being more likely to select a victim of the country with which they have the highest affinity.

Participants. Forty-three participants voluntarily took part in the study. All were students or university staff. Demographics are shown in Tables 1 and 2.

Materials. Our database contained 45 cases: all were taken from Amnesty International's web site and were active at the time of the experiment. The following values were used for the victims' and participants' attributes:

- *Age*. We used six age categories for the victims: <17, 17-25, 26-35, 36-45, 46-55 and >55. We used five categories for the participants, excluding the lowest age category, as participants had to be adults to take part.
- *Religion*. We used six religions for the victims: Christianity, Islam, Buddhism, Hinduism, Judaism, and Atheism[2]. These were chosen to cover the religions participants were most likely to have and include the four most popular world religions. For participants, we added "Other" (as the list of religions could not be exhaustive) and "I choose not to disclose" (for ethical reasons).
- *Countries*. We used twelve countries (see Table 4), chosen to (a) provide a reasonable overlap with the countries of origin of our student population (e.g. we have many students from Nigeria), and (b) have known human-right issues.
- *Case categories*. We used four case categories: death penalty, violence against women, freedom of speech, and refugees. Case categories were chosen such that there were possible cases for all the countries listed.

Procedure. The procedure was as follows:

[2] Atheism is not really a religion, but is included as a common alternative to having a religion.

- *Introduction.* Participants were told that the study will help to build a system to support people taking part in Amnesty International actions, in particular writing appeals and letters expressing concern about human rights violations for a particular victim. They were told that the study had three parts. In part 1, their preferences would be asked regarding the case they want to take action in (for example, what gender they would prefer the victim to have) so that the system can select an appropriate case from the available cases. They were told to choose even if they did not have a preference and that they would have the opportunity later to indicate how important the criteria were to them. In part 2, data about themselves would be collected. Finally, the case would be presented that matched their criteria best.
- *Part 1*. Gathering preferred attributes of the victim[3]
  - Participants selected the age, gender, country, religion and case category of the victim they preferred to take action about.
  - Participants ranked the relative importance of these criteria.
- *Part 2*. Gathering data about the participant
  - Participants selected their own age, gender, and religion.
  - Participants indicated for each country (using checkboxes) whether it was their country of origin, whether they had family living there, whether they had friends from there, and whether they had visited the country.
- *Part 3*. Participants were presented with a case that best matched their preferences. To decide on this case, the following algorithm was used. The set of cases was reduced to cases that matched the participant's most important criterion. If no such case existed, the next criterion was taken into account instead. Next, the resulting set was reduced by using the next most important criterion. This was done iteratively, till only one case remained or all criteria had been used (in which case a random case was selected from the remaining set).

Statistical Analysis. For each attribute (age, gender, religion, country), a Fisher's exact test was used to decide whether the relationship between the participants' value for that attribute and the chosen value for the victim was statistically significant[4]. If the relationship was significant, Goodman and Kruskal tau was used to determine the strength of the relationship (the extent to which the victim's attribute value could be predicted from the participant's attribute value). A Binomial test was used to decide whether significantly more participants selected a victim with the same value (for example, their own age) than predicted by random choice.

2.2 Results and Discussion

Tables 1-4 provides an overview of the results, showing (a) the proportion of participants who selected a victim's attribute in correspondence to their own attribute and (b) the expected proportion of participants who would have done this if they had

[3] We did not let participants pick victims from a list as it would have been implausible that victims existed with all attribute combinations, and the list would have been too long.

[4] Fisher's exact is used rather than Chi-square because our sample sizes are small.

made a random choice. The latter is equal to one divided by the number of answer possibilities, for example, for age 1/6, as there were 6 possible answer categories.

Does Participants' Age Influence the Selection of Victims' Age? Table 1 shows the relation between the participants' age and the victims' age. More than half the participants (63%) selected a victim of the same age, which is significantly higher than the 17% predicted by random choice (Binomial, $z=8.1$, $p<.001$).[5] Hypothesis H1.1 is supported. This seems mainly due to the youngest participants: a staggering 75% of the youngest participants (17-25, the age range of most undergraduate students) selected a victim of the same age (Binomial, $z=6.3$, $p<.001$). This is not merely due to an effect of younger victims in general provoking more support: there was a younger age category available than the one selected by these participants. However, it is noteworthy that overall 70% of the participants who had not selected a victim of the same age selected a younger victim.

Table 1. Correspondance between participants' age and victim's age

Participants'		Number	% which selected same for victim	% which would have selected same for victim if a random choice were made
Age	Total	43	63%	17%
	17-25	16	75%	
	26-35	19	63%	
	36-45	4	50%	
	46-55	2	50%	
	> 55	2	0%	

Does Participants' Gender Influence the Selection of Victims' Gender? Table 2 shows the relation between the participants' gender and the victims' gender. The relationship is statistically significant (Fisher's exact test, $p<.001$) and quite strong (Goodman and Kruskal tau=.34, $p<.001$). A clear majority of participants selected a victim of their own gender (77% compared to expected 50%, Binomial, $z=4.0$, $p<.005$). Hypothesis H1.2 is supported. This is mainly due to female participants: a staggering 95% of female participants select a female victim (Binomial test, $z=4.1$, $p<.001$), while the difference for male participants is not significant.

Table 2. Correspondance between participants' gender and victim's gender

Participants'		Number	% which selected same for victim	% which would have selected same for victim if a random choice were made
Gender	Total	43	77%	50%
	Male	22	59%	
	Female	21	95%	

[5] Fisher's exact test is not significant (p=.13), which is not remarkable given the small cell sizes, so, Goodman and Kruskal tau could not be determined.

Does Participants' Religion Influence the Selection of Victims' Religion? Fifteen participants were omitted from the analysis: seven did not disclose their religion and eight picked "other" as their religion. Table 3 shows the relation between the participants' religion and the victims' religion. The relationship is statistically significant (Fisher's exact test, p<.001) and quite strong (Goodman and Kruskal tau=.41, p<.001). A clear majority selected a victim of their own religion (64% compared to expected 17%, Binomial, z=6.8, p<.001). Hypothesis H1.3 is supported. Only Islam, Christianity and Atheism had enough cases to analyse them separately. The proportion of selecting a victim with the same religion was significantly higher than expected for both Christianity and Islam (Binomial, both p<.001), but not for Atheism.

Table 3. Correspondance between participants' religion and victim's religion

Participants'		Number	% which selected same for victim	% which would have selected same for victim if a random choice were made
Religion	Total	28	64%	17%
	Christianity	15	60%	
	Islam	6	100%	
	Buddhism	0	n/a	
	Hinduism	1	0%	
	Judaism	1	100%	
	Atheism	5	40%	

Does Participants' Affinity with Country Influence the Selection of Victims' Country? Five participants have been omitted from the analysis: three participants had no affinity with any of the countries and two had an equal highest affinity with two countries. Table 4 shows the relation between the participants' country of highest affinity and the victims' country. The relationship is statistically significant (Fisher's exact test, p<.001) and quite strong (Goodman and Kruskal tau=.44, p<.001). A clear majority of participants selected a victim of the country with which they had most affinity (61% compared to expected 8%, Binomial, z=11.6, p<.001). Hypothesis H1.4 is supported. Considering individual countries of most affinity (with the exception of Morocco for which there was not enough data), the proportion of selecting a victim from the country of most affinity was significantly higher than expected for each individual country (Binomial, p<.05). Considering the type of most affinity (country of origin, have family, have friends, have travelled there), the proportion of participants selecting a victim from the country of most affinity was significantly higher than expected for each type (Binomial, p<.005) except for the "have friends there" type (Binomial, p=.05) for which there were only few cases. The strongest effect happens for country of origin: all participants who had one of the countries as their country of origin selected a victim from that country.

Table 4. Correspondance between participants' affinity with country and victim's country

Participants'		Number	% which selected same for victim	% which would have selected same for victim if a random choice
Country of highest affinity	Total	38	61%	8%
	Country of origin	11	100%	
	Have family	12	42%	
	Have friends	5	40%	
	Have travelled	10	50%	
	China	2	100%	
	India	3	68%	
	Iran	2	100%	
	Japan	0	n/a	
	Morocco	1	100%	
	Nicaragua	0	n/a	
	Nigeria	6	68%	
	Pakistan	2	100%	
	Palestine	2	100%	
	Saudi Arabia	4	50%	
	Sri Lanka	0	n/a	
	USA	16	38%	

Case Category Selection. Table 5 shows an overview of the case category which participants selected. We explored whether there is a specific case category that is more likely to motivate people to take action. In particular, we explored whether there was any influence of gender on the case category choice, as one of the categories ("Violence against women") could have led to a gender effect. However, the results show that this is not the case. The case selection is pretty much evenly distributed.

Table 5. Case category selection differentiated by participant gender

		Case category			
		Death penalty	Freedom of speech	Refugee and asylum	Violence against women
Gender	Female	3	7	2	9
	Male	5	8	3	6
Total		8	15	5	15

Relative Importance of Criteria. Figure 1 shows participants' ranking of the criteria by importance. Almost all participants ranked either the case category or the victim's nationality as most important. Participants tended to rank the victim's religion as of little importance. There was a lot of variance between participants.

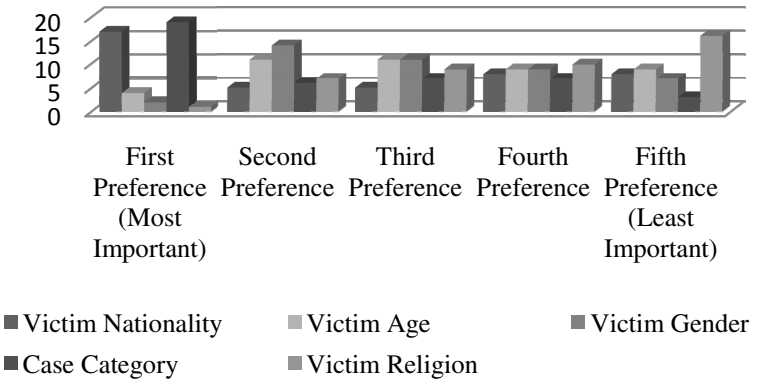


Fig. 1. Participants' perceived relative importance of criteria

3 Study 2: Kiva

In Study 1, we found that participants selected victims with similar attributes to themselves. However, a limitation of that study was that we did not check whether participants actually took more action for these victims. We wanted to verify that this bias for participants to help people who resemble them also shows in the actions people take. We therefore did an exploratory investigation of Kiva Microfunds (kiva.org): an organization which encourages ordinary people to lend relatively small amounts of money to entrepreneurs across the globe. Participants select an individual entrepreneur (or group of entrepreneurs) to lend to. They are presented with a list of people wanting loans. For each case, a picture of the entrepreneur(s) is shown, their name, country, what sector they work in (e.g. food), why they want the loan (e.g. "To buy more bananas to sell"), and a short background story. The story often mentions the entrepreneur's gender. Participants can filter cases by a.o. gender, country and sector. Participants can also become members of on-line community groups (e.g. Scotland, Christians), and there is some competition between groups about who provides the most loans. It is possible to see who the group has lent to, but this is quite cumbersome to access, so in our experience members are not very aware of the loans made by others (they are more aware of how much has been lent). So, while it cannot be excluded that members influence each other in whom they lend to, this is not that likely. Also, the religion based teams that we considered were not set up with the purpose to support people with that religion. For example, Team Christians describes itself as "A group of believers in Jesus Christ, brought together through a common purpose: to help those in need around the world."

We investigated the effects of participants' religion and gender, as it is possible to obtain the required data for this from the Kiva website. To investigate the effect of religion, we considered two Kiva community groups which clearly show their members' religion: Team Christians and Team Muslims. We considered the last 60 loans made by these teams (in March 2010). For each loan, we determined the dominant

religion in the entrepreneur's country using Wikipedia. To investigate the effect of gender, we considered two popular Kiva community groups: Team Europe and Team Obama. We considered the last 15 loans made by female members and the last 15 loans made by male members, such that: each loan was made by a different member, the member's gender was easy to identify (from their name and picture) and the loan was made to an individual entrepreneur. We found the following:

- *Religion*. 70% of Team Christians' loans were made to countries with a Christian majority (with 22% to Muslim countries, and 8% to Buddhist countries). In contrast, 85% of Team Muslims' loans were made to Muslim countries (with 12% to Christian countries, and 3% to Buddhist countries). This supports our findings from Study 1 that people help people of the same religion.
- *Gender.* In Team Europe, 93% of female participants provided their loan to a female entrepreneur, while 80% of males provided their loan to a female entrepreneur. In Team Obama, 87% of female participants provided their loan to a female entrepreneur, while 73% of male participants provided their loan to a female entrepreneur. So, most participants irrespective of their own gender provided loans to female entrepreneurs. This overall bias in providing loans to female entrepreneurs is expected given the fact that there are a lot more female entrepreneurs listed on Kiva than male ones (e.g. on 7 October 2010 there are 155995 individual cases listed as female and 51308 as male). So, the expected proportion of loans to female entrepreneurs is 75%. However, female participants still exceed this expected proportion, in correspondence with our findings in Study 1.

So, Study 2 confirms the findings of Study 1.

4 Conclusion and Future Work

In a previous exploratory study, people claimed that they do not care about the victim's attributes such as gender, religion and ethnicity when they decide to participate in Amnesty International actions. We suspected that this was due to participants being politically correct, also given the existing Psychological literature on similarity-attractiveness. This was corroborated in our first study: when forced to make a choice, participants tend to select victim attributes similar to their own. There was a significant correspondence between the participants' gender, age, and religion and those selected for the victim. Similarly, participants' affinity with countries influenced which country they selected for the victim. So, there is a clear benefit for personalizing the trigger (a case summary sent to encourage people to take part in actions) to participant attributes. Participants may be more likely to take a particular Amnesty International action when the victim's attributes (age, gender, religion) are similar to their own and when the victim's country is one they have an affinity with. These findings were confirmed in our study of participants' behaviour on micro-financing website Kiva: female participants tend to provide loans to female entrepreneurs, and participants tend to provide loans to entrepreneurs of the same religion. Nevertheless, when asked about the relative importance of factors in Study 1, participants maintain that religion is not important. This means that producing a user model by explicitly

asking users about their religion would not work, as no clear reason could be given for providing this information except that it may influence their behaviour which is unacceptable to users. On the contrary, asking users about their country of origin, countries they have travelled to and have family and friends in may be acceptable, as participants agree that the victim's country is important to them. It may be possible to infer attributes that cannot be explicitly asked, for example, a probability model of religion could be made based on the dominant religions in the country of origin. Or the user could indicate their affinity with sample countries that have a clearly dominant religion, and it could be attempted to infer their religious preferences from this.

The acceptability of personalizing to different attributes should also be taken into account in the content of the messages. For example "This case may interest you because it is about the death penalty in China" may be effective, whilst it may be better to let participants infer the gender and religion of victims (through names, referring expressions, country), unless these have an explicit bearing on the case (for example, somebody being prosecuted for their religious believes).

Based on the results, it is also possible to draw some conclusions about what personalization may be most effective. Personalizing to affinity for country seems most important, as this factor is regarded highly important by participants and is showing a very clear effect. Personalizing to gender seems more of an issue for women, so it may be possible to simply use more female than male victims (though there are some ethical problems with this, as all cases deserve support). Personalizing to age seems most effective for younger users. Personalizing to religion is difficult (as explained above), but seems quite effective for certain religions. An indirect way of obtaining information about religion, e.g. via affinity for countries, seems the best way forward.

In future, we will continue to investigate triggers in two distinct steps. First, we will investigate the best timing for delivering triggers. Secondly, we will further investigate the best content of the trigger (how much information about the victim to provide and in what form). Informed by empirical studies, we will design and develop tailoring algorithms that will select the optimal content and timing for a given user in a given situation. These algorithms will be applied in a system that will encourage people to participate in Amnesty actions, and evaluated in a longitudinal study.

References

1. Byrne, D.: Attitudes and attraction. In: Berkowitz, L. (ed.) Advances in Experimental Social Psychology, vol. 4, Academic Press, New York (1969)
2. Byrne, D.: The attraction paradigm. Academic Press, New York (1971)
3. Carenini, G., Moore, D.J.: An empirical study of the influence of user tailoring on evaluative argument effectiveness. In: IJCAI, pp. 1307–1314 (2001)
4. Fogg, B.J.: Persuasive technology: using computers to change what we think and do. Morgan Kaufmann (2003)
5. Fogg, B.J.: Creating persuasive technologies: an eight-step design process. In: Persuasive Technology Conference, Claremont, California. ACM (2009)
6. Grasso, F., Cawsey, A., Jones, R.: Dialectical argumentation to solve conflicts in advice giving: a case study in the promotion of healthy nutrition. International Journal of Human-Computer Studies 53, 1077–1115 (2000)

7. IJsselsteijn, W., de Kort, Y., 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)
8. Jackson, T.: Motivating sustainable consumption: A review of models of consumer behaviour and behavioural change. A Report to the Sustainable Development Research Network. Policy Studies Institute, London (2005)
9. Japuntich, S., Zehner, M., Smith, S., Jorenby, D., Valdez, J., Fiore, M., Baker, T., Gustafson, D.: Smoking cessation via the internet: a randomized clinical trial of an internet intervention as adjuvant treatment in a smoking cessation intervention. Nicotine and Tobacco Research 8 (suppl. 1), S59–S67 (2006)
10. Kroeze, W., Werkman, A., Brug, J.: A systematic review of randomized trials on the effectiveness of computer-tailored education on physical activity and dietary behaviours. Annals of Behavioral Medicine 31, 205–223 (2006)
11. Mazzotta, I., de Rosis, F., Carofiglio, V.: PORTIA: a user-adapted persuasion system in the healthy eating domain. IEEE Intelligent Systems 22, 42–51 (2007)
12. Michie, S., Johnston, M., Francis, J., Hardeman, W., Eccles, M.: From theory to intervention: mapping theoretically derived behavioural determinants to behaviour change techniques. Applied Psychology: An International Review 57, 660–680 (2008)
13. Midden, C., Ham, J.: Using negative and positive social feedback from a robotic agent to save energy. In: Persuasive Technology Conference, Claremont, CA. ACM (2009)
14. Nguyen, H., Masthoff, J.: Designing empathic computers: The effect of multimodal empathic feedback using animated agent. In: Persuasive Technology, Claremont, CA. ACM (2009)
15. Noar, S.M., Benac, C.N., Harris, M.S.: Does tailoring matter? Meta-analytic review of tailored print health behavior change interventions. Psych. Bulletin 133, 673–693 (2007)
16. Noar, S.M., Harrington, N.G., Aldrich, R.S.: The role of message tailoring in the development of persuasive health communication messages. In: Beck, C. (ed.) Communication Yearbook 33, pp. 73–133. Routledge (2009)
17. O'Keefe, D.J.: Persuasion: Theory and research. Sage, Newbury Park (1990)
18. Reiter, E., Robertson, R., Osman, L.: Lessons from a failure: Generating tailored smoking cessation letters. Artificial Intelligence 144, 41–58 (2003)
19. Revere, D., Dunbar, P.J.: Review of computer-generated outpatient health behaviour interventions: Clinical encounters "in absentia". Journal of the American Medical Informatics Association 8, 62–79 (2001)
20. Langrial, S., Oinas-Kukkonen, H.: Less Fizzy Drinks: A Multi-method Study of Persuasive Reminders. In: Bang, M., Ragnemalm, E.L. (eds.) PERSUASIVE 2012. LNCS, vol. 7284, pp. 256–261. Springer, Heidelberg (2012)
21. Forbes, P.J., Wells, S., Masthoff, J., Nguyen, H.: SUPERHUB: Integrating behaviour change theories into a sustainable urban-mobility platform. In: Using Technology to Facilitate Behaviour Change and Support Healthy, Sustainable Living Workshop at BHCI (2012)
22. Nguyen, H., Masthoff, J.: Is it Me or Is it what I say? Source Image and Persuasion. In: de Kort, Y.A.W., IJsselsteijn, W.A., Midden, C., Eggen, B., Fogg, B.J. (eds.) PERSUASIVE 2007. LNCS, vol. 4744, pp. 231–242. Springer, Heidelberg (2007)
23. Oinas-Kukkonen, H.: A foundation for the study of behavior change support systems. Persuasive Ubiquitous Computing, 1–13 (2012)
24. Nguyen, H., Masthoff, J.: Mary: A personalised virtual health trainer. Adjunct Proceedings of UMAP 2010, pp. 58–60 (2010)
25. Grasso, F., Ham, J., Masthoff, J.: User Models for Motivational Systems. In: Ardissono, L., Kuflik, T. (eds.) UMAP Workshops 2011. LNCS, vol. 7138, pp. 335–341. Springer, Heidelberg (2012)

Examining the Efficacy of a Persuasive Technology Package in Reducing Texting and Driving Behavior

Brenda Miranda[1], Chimwemwe Jere[2], Olayan Alharbi[2], Sri Lakshmi[2], Yasser Khouja[2], and Samir Chatterjee[2]

[1] School of Behavioral and Organizational Sciences,
[2] School of Information Systems and Technology, Claremont Graduate University,
130 E. 9th Street, Claremont, CA, 91711, USA

Abstract. Over the past decade, texting and driving has become a prevalent form of distracted driving and resulted in an alarming rate of deaths and injuries. Research has documented the debilitating cognitive effects of engaging in texting and driving, comparing it to the dangers of driving drunk. Several states have implemented legislation banning texting and driving, however it remains a national epidemic. There is a paucity of empirical research examining the effectiveness of strategies in decreasing texting and driving behavior. Research employing technology as a potential solution has focused on using strategies such as hands-free technology or monitoring devices and applications. The current study takes on a different approach, by examining the efficacy of a persuasive technology package in motivating and facilitating behavior change. Findings provide preliminary evidence for the efficacy of pairing a video documentary and text message reminders in decreasing texting and driving behavior.

1 Introduction

In previous decades, drivers often engaged in more traditional distracting activities like drinking or eating while driving; however, as mobile technologies have evolved such as cell phones, and GPS devices, more nuanced distraction sources are readily available. Distracted driving plays a pivotal role in traffic injury and fatalities [1]. In 2010 alone, over 3,000 drivers were killed in accidents related to distracted driving and an additional 416,000 were injured [1].

As text messaging has become increasingly common in United States and worldwide, it has become one of the most prevalent forms of distracted driving. Cell phone use alone accounts for 2,600 vehicle fatalities and 300,000 collisions annually [1]. A recent study found that 1 in 4 American adults have either sent or read text messages while driving, and that young adults are as likely as teens to text while driving [3]. Research indicated that texting while driving makes one 23 times more likely to crash than a non-distracted driver [1].

Studies suggest that texting and driving is riskier than talking on the cell phone or to fellow passengers. When texting, drivers respond slowly to brake lights in front of them and show impairment in forward and lateral control [4]. Texting requires

S. Berkovsky and J. Freyne (Eds.): PERSUASIVE 2013, LNCS 7822, pp. 137–148, 2013.

drivers to switch their attention from one task to other, and thus their overall reaction times are substantially slower than when they are engaged in phone conversation. Previous studies have found that texting while driving results in driving that is as or more debilitating than driving legally drunk (i.e., BAC at .08%) [6]. Interestingly, drivers often underestimate the degree of distraction associated with texting and driving, and feel that they can adequately drive while texting [5].

Although some acknowledge that texting and driving is risky and regardless of the fact that texting and driving is illegal in some states, many individuals continue to engage in this behavior. A study examining some of the motivating factors behind cell phone use while driving, found that people are likely to engage in this behavior if they perceive the conversation they are having is important and believe they are good at multitasking [5]. According to this study, perceived importance of the conversation was a higher predictor of cell phone use while driving, than perceived risk.

Although texting and driving is a growing epidemic, there is little empirical work examining strategies for reducing this behavior. Currently 39 states, the District of Columbia, and some local governments prohibit all drivers from texting and driving [1]. Studies examining the consequences of establishing bans against texting and driving have mixed results. While some studies have found that after implementing laws banning texting and driving reduces personal injury accident rates, other studies find that the rates of cell phone usage actually increase [7]. In recent years, several public service announcements (PSA) have been developed to discourage texting and driving. In doing so, several PSAs have focused on utilizing fear appeals to convey their message. A recent study examined the use of fear appeals in discouraging texting and driving behavior, and found that after viewing two fear appeals participants reported viewing texting and driving behaviors as more distracting than previously believed, but also reported an increased intention to engage in texting and driving behavior [7]. This behavior phenomenon is referred to as the boomerang effect and is believed to occur as a result of participants' reaction to the message and denial of a perceived threat [7]. The current study examines the efficacy of a persuasive technology package in decreasing texting and driving behavior, by motivating and facilitating behavior change.

2 Theoretical Framework

BJ Fogg's behavioral model proposes that three elements must be present for a behavior to occur: motivation, ability, and a trigger [8]. The present study acknowledged the fact that while individuals are *able* to refrain from driving, they are often *unmotivated* and have no *reminder to facilitate* their behavior change. Fogg argues that a behavior does not occur when one or more of these elements are missing. Triggers play a particularly interesting role in the equation, as they stimulate awareness and can in some cases promote increases in the other two elements, motivation and ability. However, triggers themselves may not be enough to convince individuals to accept a persuasive message and change their behavior. The current study was designed to decrease texting and driving behavior using a persuasive technology package that included elements to serve as a trigger for behavior change, but also elements to increase participants' motivation to want to change their behavior.

2.1 Persuasive Technology Package: Technology as Media and Tool

Fogg defined the persuasive technology as "a computing system, device, or application intentionally designed to change a person's attitude or behavior in a predetermined way" [8]. His framework for the roles technology play in persuading others, the functional triad, identifies three ways persuasive technology can operate: as tools, as media, and as social actors [8]. In the current persuasive technology package, technology functioned as media and tool.

The media element was a 10-minute documentary video produced by AT&T [9], and edited by researchers to a shortened version of 5 minute and 22 seconds. As BJ Fogg describes, media can often serve the purpose of providing vicarious experiences that motivate individuals to partake in a particular behavior [8]. According to Fogg's principle of cause and effect, media can persuade people to change their attitudes or behaviors by enabling them to observe immediately the link between cause and effect without experiencing the risk of consequences. The documentary was included to expose participants to real-life cases of both victims and perpetrators of accidents caused by texting and driving behavior.

The tool element was comprised of a daily text message sent in the morning across three days. B.J. Fogg describes a persuasive technology tool as an interactive product designed to change attitudes or behaviors or both by making desired outcomes easier to achieve. The text message reminder was designed to interact with participants before their morning commute, and extend the effects the video documentary may have on participants' attitudes. The interactive component involved sending messages to each participant and prompting each participant to reply to this message.

3 Methodology and Research Design

The current study employed the use of persuasive technology, by pairing a video and one daily reminder text message sent across three days, to decrease texting and driving behavior. The efficacy of the current persuasive technology package was examined using a repeated measures design to measure change in participants' behavior over time. A repeated measures design involves data collection of participant behavior at two different time points. This design allowed for a comparison of participants' texting and driving behavior before and after experiencing the current persuasive technology package.

3.1 Participants

Participants were 37 individuals, 20 males and 17 females, with a mean age of 21.6 years and an age range of 21-34 years, recruited using convenience sampling. Convenience sampling is characterized by the selection of participants based on their availability and proximity to researchers. The current study was to examine the efficacy of combining a video and text-messaging platform in persuading individuals to decrease texting and driving behavior, thus the focus was not on generalizing these results to the general population, but rather to generate preliminary evidence that the

current use of persuasive technology is effective in achieving that goal. Participants were recruited based on two criteria a) current participation in texting and driving behavior and b) aged under 35.

3.2 Video Component as Persuasive Media

The 5-minute video documentary served as the media component, by providing participants the ability to vicariously explore the cause and effect relationships of texting and driving behavior. This video depicted two factual stories regarding the dangers of texting while driving. Each story presented two different perspectives. The first was a story about a teenage girl who lost her own life while texting and driving, a few days before her high school graduation. The second was about a young man who while texting and driving, ran over a bicyclist who ended losing his life. Both stories emphasized the fact that both individuals frequently sent text messages and had become near-experts in multitasking. The video had two main goals. The first was meant to emphasize that although individuals text and drive frequently and feel capable of doing so, does not mean they are not susceptible to accidents as a result of this behavior. The video also aimed to connect the potential consequences of texting and driving to real stories that would give these consequences more meaning. Our video aimed to bring two factual experiences together, to persuade participants to pledge to stop texting and driving and agree to receive a daily text message for three days. The edited version of the AT&T video that was used in this project can be watched at http://goo.gl/QWwcD.

3.3 SMS Text Component as Persuasive Tool

Text message reminders sent served as a tool to make it easier for participants to stop texting and driving. More specifically, these messages served as suggestion technology, to remind them to refrain from texting and driving. These messages were designed to *trigger* the avoidance of texting and driving at times when this behavior occurs. The timing of suggestion is critical [8], thus messages were sent before participants presumably began their morning commute, at 7 am. Participants were in different time zones, thus message sent-time was tailored for participants in each time zone.

Participants received personalized messages that began by addressing each participant by their first name. Each message attempted to achieve the goal of decreasing texting and driving behavior in a different manner (e.g., stating a fact, encouraging participants to rethink whether a text is worth hurting themselves or others, presenting the avoidance of texting and driving as a social norm). All messages ended with the following prompt: "Do you pledge to avoid/not text and drive today? Y/N." For instance, a sample text message may read, "Brenda, 5sec = the avg time ur eyes are off the road while texting @ 55mph that's equivalent to driving over a whole football field. Do u pledge not to text&drive today? Y/N."

To facilitate the delivery of multiple text messages to participants in various time zones, text messages were sent via an affordable cloud-based Short Message Service (SMS) text-messaging platform. The Twilio platform with built-in API libraries

supports multiple programming languages (e.g. PHP, Python, .NET, Perl, Java, C++). Twilio's cloud powered SMS Application Programing Interface (API) design is a service with outbound and inbound SMS charge rate of 0.01 USD per 160-character message. After downloading Twilio PHP API libraries, a PHP script-based web application was setup to send and receive mobile text messages via Twilio. SMS analytics, including messages and notifications, were logged and tracked under Twilio's comprehensive dashboard.

3.4 Procedures

Participants were prompted via email to watch a 5-minute video in the form of a mini-documentary depicting two factual cases related to the consequences of texting while driving. Participants then had the option of taking a pledge to stop texting and driving and receive a daily text message over the next 3 days. Upon agreement, participants were redirected to a 5- item electronic survey regarding their current texting while driving behaviors and their experience with the video (Appendix A). A novel text message reminder was sent each morning (at 7 am) during the following 3 days. Each text message was designed to address a different goal: a) prompt participants to compare the relative importance of the text in comparison to their own safety and the safety of others, b) provide a novel fact about texting while driving and c) encourage participants to see the avoidance of texting and driving as a social norm. After receiving all daily text messages, participants were sent a 7-item electronic survey regarding the text messages they received and their texting and driving behavior after taking part in the current study (Appendix B).

4 Results

Based on replies to the daily text message reminders, throughout the 3-day intervention period, all participants (100%) pledged to stop texting while driving for a single day, showing the level of commitment of participants to stop texting. Each pledge was instantly acknowledged in order to enhance interactivity between the persuasive system and the participants, and motivate the latter's continued participation (i.e., <John Smith>, Thanks for participating today! Powered by: Stop-Texting-n-Driving Team@CGU).

Table 1 describes participants' experience in watching AT&T's mini-documentary. The post-video survey shows that the majority of participants agreed to some degree that the on-line video persuaded them to rethink their ability to text while driving (91.9%). Furthermore, the majority of participants agreed to some degree that the on-line video brought to life some of the consequences of texting while driving (91.9%). Participants were also asked to rate the effectiveness of the video in persuading them to stop texting and driving (Table 2). The survey results indicate that the majority of participants rated the on-line video as an effective tool to persuade them to stop texting while driving (81%). According to the post-video survey results, viewing of the on-line video invoked varying emotional feelings among the participants, with the majority feeling some form of sadness (54.1%).

Table 1. Percentage of participant agreement with the predetermined goals of the video

The video….	Strongly Disagree	Disagree	Neither Agree nor Disagree	Agree	Strongly Agree
Made me rethink my own ability to text and drive	2.7%	0.0%	5.4%	54.1%	37.8%
Brought to life some of the consequences of texting and driving	0.0%	5.4%	2.7%	29.7%	62.2%

Table 2. Participants' ratings of the video's effectiveness in persuading them to stop texting and driving

	Very Ineffective	Ineffective	Neither Effective nor Ineffective	Effective	Very Effective
How would you rate the overall effectiveness of the video in persuading you to stop texting and driving?	2.7%	0.0%	5.4%	54.1%	37.8%

As part of the post-video survey, participants were requested to honestly report the average number of times they text while driving per day (Figure 1). Before participating in the current study, participants reported texting and driving on average between 0 and 17 times per day. Notably, although we recruited participants based on predetermined knowledge that participants engaged in texting and driving behavior, 27% indicated that they do not text while driving at all. Only participants who reported engaging in texting while driving were included in comparative pre- and post-intervention analyses.

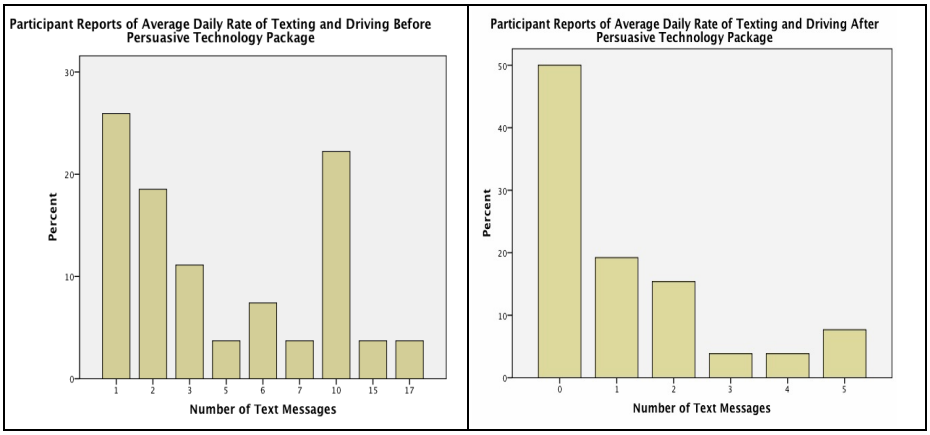


Fig. 1. Participant reports of daily texting while driving rates pre- and post-persuasive technology package

After exposure to the persuasive technology package, of the participants that reported engaging in texting and driving behavior, participants wrote between 0 and 5 text messages per day while driving (Figure 1). To determine whether there were statistically significant differences between participants' reports of average daily texting and driving behavior before and after exposure to the current persuasive technology package, a Wilcoxon T test was conducted. This nonparametric test allowed the examination of participants' reports at both time points, while having a small data set and without the need to satisfy the assumption of normal distribution. Only participants who reported engaging in some texting and driving behavior in the pre-survey ($n = 26$) were included in this analysis. Findings revealed that on average participants significantly decreased their daily rate of texting and driving after exposure to the current persuasive technology package ($Z = -4.306$, $p < .001$). Before exposure to the current persuasive technology package, participants reported an average of 5.26 text messages while driving, and after participating they reported an average of 1.57 text messages while driving. Of the participants included in this analysis, 2 remained with no change, and 24 decreased the average daily texting and driving rate to some degree.

In the post-survey, participants had the opportunity to rate their experience with the text message reminders compared to the predetermined goals of these text messages (Table 3). Generally, the majority of participants agreed to some degree that the text messages were a good reminder to avoid texting and driving (82.8%), increased their motivation to stop texting and driving (74.3%), and increased their ability to stop texting and driving (62.8%). Participants also rated which daily text message they felt was most effective. Most participants (40%) felt that the fact-based text message was most effective followed by the text comparing the relative importance of sending a text while driving, to the safety of themselves and others (37.1%). Lastly, although some participants reported not engaging in texting and driving behavior, all participants were asked to what degree their texting and driving behavior had changed as a result of their participation in the study (Table 4). Most participants reported decreasing their texting and driving behavior (60%) or completely stopping texting and driving (34.3%).

Table 3. Participants' rating of their experience with the text message reminders compared to the text message's predetermined goals

The daily text messages:	**Strongly Disagree**	**Disagree**	**Neither Agree nor Disagree**	**Agree**	**Strongly Agree**
Were a good reminder to stop texting and driving	0%	8.6%	8.6%	45.7%	37.1%
Increased my motivation to stop texting and driving	0%	14.3%	11.4%	40%	34.3%
Increased my ability to stop texting and driving	8.6%	20%	8.6%	45.7%	17.1%

Table 4. Participants' report of change in texting and driving behavior as a result of participating in the current study

As a result of my participation, I have….	**Percent**
Completely stopped texting and driving	34.3%
Decrease my texting and driving	21%
Continued texting and driving at the same rate	5.7%
Increased my texting and driving	0%

5 Discussion

Distracted driving has become a nationwide epidemic, affecting drivers on roads across the country. A particular emphasis has been placed on texting and driving behavior, as technology, particularly cell phones with texting-capacity have become increasingly prevalent. Recent research has found that drivers who are texting are 23 times more likely to crash than drivers who are not distracted [1]. Given the magnitude of this issue, more research must be conducted to measure the efficacy of ways to diminish this dangerous behavior. Based on the current findings, there is evidence that the employed persuasive technology package is effective in decreasing individuals' texting and driving behavior. We found statistical evidence that there was a significant difference (for participants who reported engaging in texting and driving behavior) in average daily reports of texting and driving behavior before and after the current study.

Interestingly, even some participants who initially reported not engaging in texting and driving behavior reported some decrease in texting and driving behavior. This is evidence that participants' responses may have been affected by social desirability bias [11], particularly after watching a video depicting the consequences of their actions. More specifically, because respondents want to present themselves in the best possible manner and have a tendency to respond in a "socially acceptable manner," participants may have been reluctant to admit they engage in a behavior as dangerous as texting and driving, particularly after viewing real cases of individuals who have been gravely affected by this behavior. However, out of a sample of 37 participants, 27 of which reported some engagement in texting and driving behavior, all but 2 participants (i.e., 35) reported that they decreased or completely stopped texting and driving after participating in the current study. One inherent weakness of the current study design is the participants' subjective rating of their own texting and driving behavior. As such, the study is prone to validity problems because participants, in some cases, may exaggerate or under-report the severity of their texting and driving behavior. Future research should include multiple measures of participant behavior that can address the bias of self-report.

From an effective persuasive systems design perspective, our study employed a combination of two forms of persuasive technology (i.e., video media and text messages) to achieve the various goals of our project. The video was designed to provide a vicarious experience of the consequences of texting and driving, as well as invoke an emotional appeal that would allow viewers to sympathize with the two

factual cases presented. To connect this visual and emotional experience to participants' daily lives, the current study employed unique actionable daily reminder SMS text messages. These messages not only reminded participants to avoid texting and driving, but also allowed participants to actively reply to the text message with a daily commitment to avoid texting and driving.

6 Future Studies

In this study, we designed and employed a persuasive technology package based on a video and reminder SMS text messages that aimed to reduce texting while driving behavior. While the current study found evidence of the efficacy of the current intervention in decreasing texting and driving behavior, a number of limitations can be addressed in future research. Although the current study attempted to send text message reminders at an opportune time (before participants' morning commute) to be able to affect participant behavior throughout the day, participants domiciled in different time zones and have different work schedules. From participant feedback, we determined that some participants were receiving text messages too early or too late based on their specific morning commutes. Future studies should therefore explore ways to easily sync SMS text message to each individual's commuting time, based on ideal times identified by each participant.

Participants generally liked the on-line video and found it profoundly compelling because they remembered it every time they were tempted to text while driving. Such participants, after viewing the on-line video, took the initiative to recruit other participants in order to share their profound experiences. In light of the positive findings of the current study, future studies can explore the current intervention in a large-scale study that distributes the video to a wider audience and also allows participants to share the video with others (e.g., via link in text, email, or social network). The current persuasive technology package linked all three steps of the intervention via a one-click redirection, making it possible for any individual to watch the video, sign up for text message reminders, and take a survey about their texting and driving behavior through a sequence of clicks. Future studies should consider the use of this format because it encourages participants to fulfill all steps of the study and decreases attrition.

Lastly, given our observations of the potential effects of social desirability bias, future studies should consider surveying participants' pre-intervention rate of texting and driving behavior before engaging in any part of the study [e.g., watching a video], as participants may be more likely to respond candidly if they have not been exposed to any factual or persuasive information regarding texting and driving. Future studies may also include alternate ways of measuring participant behavior that reduce the bias of self-report.

7 Conclusion

In this study, we explored how to use a combination of a media tool with mobile SMS technology to persuade participants to reduce or stop texting while driving behavior.

The combination of the two forms of persuasive technology was effective in achieving a short-term behavior reduction or termination of texting while driving behavior. A survey follow-up to the current study will be conducted three months post-study, to determine the long-term effects of the current persuasive technology package on participants' texting and driving behavior. Findings yield preliminary evidence that the current persuasive technology package is a promising avenue for reducing texting while driving behavior in young adults (20 – 35 years).

We employed a mobile SMS text messaging platform as a channel for communicating persuasive messages to participants. As a persuasive system, mobile SMS technology offers many advantages over traditional media including scalability, greater anonymity, interactivity, ubiquity, as well as handling of high volumes of data.

References

1. NHTSA: National Highway Traffic Safety Administration, http://www.distraction.gov/
2. Cook, J., Jones, R.M.: Texting and Accessing the Web While Driving: Traffic Citation and Crashes Among Young Adult Drivers, pp. 545–549 (2011)
3. Benden, M.E., Smith, M.L., Henry, M., Congleton, J.J.: Reviewing Four Decades of Cell Phone Use while Driving Literature (1970-2010): An Emphasis on Texting Behaviors, Parental Perceptions, and Methods of Control. Journal of Health Behavior and Public Health 2(2), 20–26 (2012)
4. Drews, F.A., Yazdani, H., Godfrey, C.N., Cooper, J.N., Strayer, D.L.: Text Messaging During Simulated Driving (2009)
5. Nelson, E., Atchley, P., Little, T.D.: The effects of perception of risk and importance of answering and initiating a cellular phone call while driving. Accident Analysis & Prevention 41(3), 438–444 (2009)
6. Strayer, D.L., Drews, F.A., Crouch, D.J.: A comparison of the cell phone driver and the drunk driver. Human Factors 48(2), 381–391 (2006)
7. Lennon, R., Rentfro, R., O'Leary, B.: Social marketing and distracted driving behaviors among young adults: The effectiveness of fear appeals. Academy of Marketing Studies Journal 14(2), 95–104 (2010)
8. Fogg, B.J.: Persuasive Technology. Morgan Kaumann, San Francisco (2003)
9. AT&T: Don't Text While Driving Documentary, https://www.youtube.com/watch?v=DebhWD6ljZs
10. Oinas-Kukkonen, H., Harjumaa, M.: Persuasive Systems Design: Key Issues, Process Model, and System Features. Communications of the Association for Information Systems 24 (2009)
11. Maccoby, E.E., Maccoby, N.: The interview: A tool of social science. Handbook of Social Psychology 1, 449–487 (1954)

Appendix A. Pre-survey

Please fill out the following questions about yourself and the text messages you received. All information you provide will be completely confidential. Please respond as accurately as possible.

Rate your level of agreement with the following statements regarding the video you watched.

The video…

	Strongly Disagree	Disagree	Neither Agree nor Disagree	Agree	Strongly Agree
1. Made me rethink my own ability to text and drive	○	○	○	○	○
2. Brought to life some of the consequences of texting and driving	○	○	○	○	○

2. After watching the video, I felt (insert emotion)

3. How would you rate the overall effectiveness of the video in persuading you to stop texting and driving?

- Very Ineffective
- Ineffective
- Neither Ineffective nor Effective
- Effective
- Very Effective

4. It is CRITICAL to this project that you respond HONESTLY. Your responses will be kept COMPLETELY CONFIDENTIAL. On average, how many times do you text and drive per day?

Appendix B. Post-survey

Please fill out the following questions about yourself and the text messages you received. All information you provide will be completely confidential. Please respond as accurately as possible.

Rate your level of agreement with the following statements regarding the daily text messages you received.

The daily text messages…

	Strongly Disagree	Disagree	Neither Agree nor Disagree	Agree	Strongly Agree
1. Were a good reminder to avoid texting and driving	○	○	○	○	○
2. Increased my motivation to stop texting and driving	○	○	○	○	○
3. Increased my ability to stop texting and driving	○	○	○	○	○

2. Which text message did you find most effective?

- Is that text worth hurting yourself or others? While driving today, put your phone on silent. Do you pledge to avoid texting&driving today?
- 5sec = the avg time ur eyes are off the road while texting @ 55mph that's equivalent to driving over a whole football field. Do u pledge not to text&drive today?
- At school/work u silence your phone out of respect. Respect your own & others' safety by silencing your phone when u drive. Do u pledge not to text&drive today

3. It is critical to this project that you respond *honestly*. Your responses are completely confidential. On days you received text message reminders, on average, how many times did you text and drive per day?

4. As a result of my participation, I have...

- Completely stopped texting and driving
- Decreased my texting and driving
- Continued texting and driving at the same rate
- Increased my texting and driving

5. Please provide any feedback related to the video or text messages received. Your comments are valuable to us.

Enhanced Reach: Assisting Social Interaction Based on Geometric Relationships

Asaki Miura, Dushyantha Jayatilake, and Kenji Suzuki

Artificial Intelligence Laboratory
University of Tsukuba, Japan
asaki@ai.iit.tsukuba.ac.jp, {dush,kenji}@ieee.org

Abstract. Social interaction among children plays a significant role in their social development. Some children, however, find it difficult to initiate interaction and there are only few tools that can create opportunities for children to interact with others.

This study presents a small wireless device that can measure and visualize geometric relationships in a gymnasium or playground. The estimation of geometric relationships is proposed based on signal strength of wireless communication, bodily orientation and statistical geometric consistency. A light-emitting visualization method is used in real-time according to geometric relationships among devices. Several wearable interfaces were developed to facilitate communication and social interaction of children by using the developed wireless device. Several experiments were done with typically developing children and children with pervasive developmental disorders (PDD) to evaluate the proposed technology.

1 Introduction

Social interaction among children plays a significant role in their social development; and interaction with peers helps children's learning and development [1]. Children usually love to play and interact with others. However, due to various medical and psychological conditions such as PDD, some children find it difficult to initiate interaction and only a limited number of tools are available to assist them [2]. We aspire to develop tools that can primarily create opportunities so that children could interact with others voluntarily. One of the best places to support such a interaction, in particular among children, is the playground. We have been developing playware that aims to produce play and playful experiences among users [3], through devices integrated with intelligent hardware and software. Specifically, we focus our investigation on social playware that can communicate between each other to promote social interaction.

A group is a social relationship where two or more persons interact with one another in such a manner that each person influences and is influenced by each other [4]. Group Dynamics has been studied since many years in several disciplines including children and play; and promoting social interaction would be influenced by observing and guiding group dynamics.

S. Berkovsky and J. Freyne (Eds.): PERSUASIVE 2013, LNCS 7822, pp. 149–154, 2013.

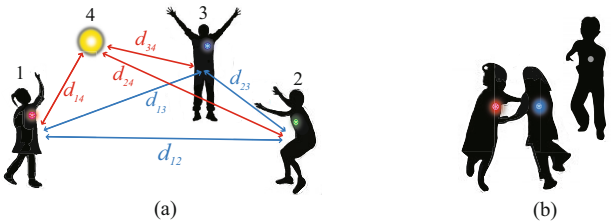


Fig. 1. Enhanced Reach: Inviting you to play with others. (a) change of ball color indicating where to pass it next (b) A kid from outside joining a group after seeing the badges illuminate.

This paper presents a novel work on a lightweight, wireless device called ER (Enhanced Reach), which not only observes and influence group dynamics, but also facilitates social interaction of children. It is developed as a *social playware*, which is designed to support and improve social interaction. In contrast to typical sports equipment and assistive tools, social playware aims to naturally generate interest or encourage approach towards the others.

Physical position relation, body postures, and actions are results of social behavior of people. Getting closer, separating, and facing each other can be seen at every scene in daily life. Functional aspects such as turning towards the partner or changing body posture to represents one's interest are indicators of social relationship, and these are significantly related to cognition and emotions of people. If a particular playware could measure positions of each other as illustrated in Fig. 1, we consider that it is possible to encourage people to interact with them and eventually create opportunities to initiate social interaction by appropriately making visualizations based on those spacial positions.

2 ER: Enhanced Reach

ER is a small wireless device consisting of position sensing, wireless communication and visualizations, and could be embed into sportswear or play tools.

Estimation of Geometric Relationships: There are various different ways to obtain geometric relationships of devices [5]. However, as wireless communication is an essential component to exchange information, we have decided to use the received signal strength indicator (RSSI) of communication between devices, where the distance is inversely proportional to the RSSI. However, with ER embedded to sportswear, when the players are not facing each other, their bodies could become obstacles to RSSI. Therefore, we considered that when the signal strength between devices i and j are I_{ij} and I_{ji}, and their orientations are θ_i and θ_j, the distance $d_{ij}(t)$ between i and j at time t could be expressed from function F as:

$$d_{ij}(t) = F(I_{ij}(t), I_{ji}(t), \theta_i(t), \theta_j(t)) \tag{1}$$

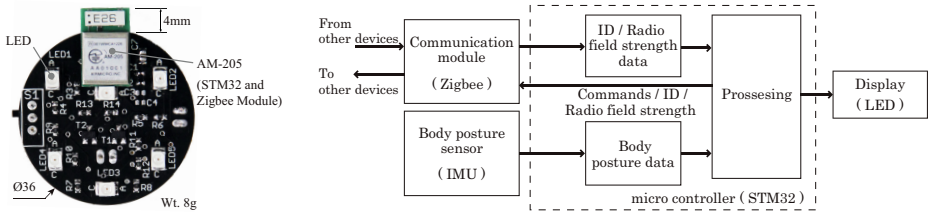


Fig. 2. The ER device and the System Configuration

Fig. 3. 3 different types of wearable ER devices: (a) badge, (b) jersey and (c) bracelet

Afterwords, we used non-metric multidimensional scaling (MDS) of Sammon [6] to convert these relative distances to 2D coordinate positions of devices and obtain their geometric relationships. Finally, the geometric relationships were used to determine visualizations. Mainly three different modes were used for visualization: (i) Illuminate the devices when close, (ii) Illuminate the devices when more than n devices are close and (iii) pass the light to the neighbour.

System Overview: Fig. 2 shows the overview of the ER device. For communication and control, we used the ZigBee AM-205 module (AIRMICRO Inc.), which consist of a chip antenna and an STM32W microcontroller (STMicroelectronics Inc.). The Inertia Measurement Unit (IMU) was used to determine orientation and LEDs were used for the display. Obtaining geometric relationships and determining LED status were performed in the processing unit.

Wireless Communication Protocol: The wireless module communicates IDs and RSSI among the devices present in the network. Considering real-time communication, we used a star type network topology with one parent device and multiple child devices. The advantage with this topology is that the parent could also perform as a regular ER simultaneously.

At the beginning, when the send command is issued by the parent, the children broadcast their IDs, so that both the parent and children in the network can measure their RSSI with each other. Afterwords, when a send command is issued by the parent, children first send their IDs, orientations and collected RSSI from other modules to the parent followed by broadcasting their IDs again. The network RSSI could be recalculated for the next step. This system can currently operate at 26 acquisitions per second with up to 10 devices.

Fig. 4. The map of event location and children collaboratively using their ER bracelets to find the suspect with the ER badge

Wearable Interfaces with ER: Fig. 3 shows 3 different wearable devices developed by embedding ER. The badge type is small and can be used easily, for instance, by affixing to clothing. The jersey type is suitable for sports exercise or training. The ER module is near the chest area, and LEDs are placed around the collar for easy visualization. The bracelet type is also developed and it is easy to see your own visualization.

3 Evaluation Experiments

We empirically obtained the simplified function for F in eq. 1 with $\alpha = 11.8$ and $\beta = -4.03$ by conducting field tests in outdoor and gymnasium environments. The RSSI was set in the range of $-98\,\mathrm{dbm} < I_{ij} < 0\,\mathrm{dbm}$. For simplicity, we did not use IMU data in this experiment.

$$F(I_{ij}) = \exp(-\frac{I_{ij} + \beta}{\alpha}) \tag{2}$$

ER Bracelet: We participated in a public social activity event for children ("Good Job event", Tsukuba city). It is an event for children up to primary school level to experience role-playing of important characters in everyday life. We put-up a police booth and developed a game for children to act as police and find a suspect in a crowded ground (Fig. 4, left). Each child in the group was given an ER-bracelet to help guide him/her to the suspect who was wearing an ER-badge. For this game, we used the visualization conditions of no illumination when distance between devices is more than 20 m, blinking for less than 20 m and always on for less than 2 m.

About 130 children participated in groups of about 10 children at a time, and most of them did not know each other. First, children started to search individually while only checking the light of their own ER. But they gradually started to cooperate and communicate with others, and started to depend on readings from multiple ER's to search for the common target and to cooperatively catch the target. Fig. 4 shows some of the reactions of the children. This game confirms that the ER could be used as a social playware and create opportunities for interaction.

ER Jersey and Evaluation of Group Dynamics: 14 children (2 typically developing, 12 with PDD) between ages 6 and 13 have participated in the experiments in which weekly sessions were conducted over a period of 3 months.

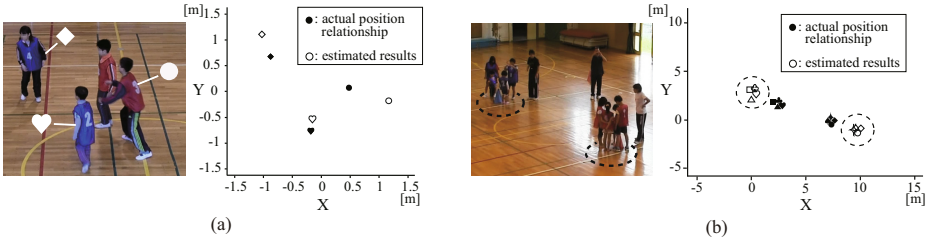


Fig. 5. Evaluation of group dynamics. (a) actual and estimated positions of 3 children in closed proximity, (b) children moving in 2 tight groups.

Fig. 6. Reaction of children to response of ER

The experiment was conducted in collaboration with a non-profit organization that supports social interaction of children with developmental disorders through the organizing of regular physical activities like sport games for them. The informed consent is obtained from the parent with the aid of the organization. Evaluation was 2 fold: quantitative analysis of group dynamics and qualitative analysis of interaction.

Two group games; a basketball game and a game to find the members who light together were performed. The basketball game was performed without using LED visualizations for the purpose of evaluating the measuring of group dynamics. Fig. 5 shows the evaluation of group dynamics with ER. Actual positions were obtained through off-line analysis of video recordings. In the member finding game, while using the visualization pattern illuminate when close, ER-jerseys were prepared by pre-programming their device IDs into 2 groups. When children first joined our experiment, they did not know that the visualizations in ER are based on geometric relationships. However, once the children felt that the jerseys light when they get closer, they tried to confirm it by getting closer, making groups and inviting others. Once they got accustomed to the concept, we observed them even holding hands with each other (Fig. 6). Furthermore, we observed them voluntarily moving towards others, calling others over even by their names and getting absolutely delighted in discovering their group partners. It was great to see them paying attention to others and cooperating with others.

4 Discussion and Conclusion

The good job event was held at a large, outdoor environment over a period of 2 days with over 130 participants and our devices were handled by a large number

of children; confirming the performances and robustness of the devices. When the first blinking of a device occurred, other children ran to that child to see it blinking as well as to have their own devices lit; and started to discuss where the suspect could be. With ER we managed to create an opportunity for children to initiate communication and make new friends.

As illustrated in Fig. 5, ER can quantitatively obtain group dynamics, whether the children are moving separately or moving around in groups. As a result, it was possible to quantify the amount of interaction such as how often children get together or who stays behind; providing valuable information for the game assistants (therapist) to focus their attention. In the group forming game, the 2 groups appears to be further apart than they really are (Fig. 5(b)) due to obscuring from their bodies. However, main focus in this study is about individual child's behaviour as a group member, and we can clearly distinguish who are in the group and who are not as seen in Fig. 5(b). Nevertheless, we are working to improve this error due to body posture by incorporating IMU data.

This paper presented the development of a new small wireless device: the Enhanced Reach (ER) to measure and visualize geometric relationships of people. Three types of social playware were developed to promote and facilitate communication of children. Experiments were performed to evaluate the reliability of proposed technology; and the usability of playware were verified by testing them with children with PDD and in an open event for children held at the Tsukuba city. We made it possible to develop games to assist and evaluate communication, and obtained geometric relationship among children during the whole game. In the future, we plan to visualize group dynamics on a tablet device so that game styles and LED visualizations could be modified in real time by trainers or therapists.

References

1. Howe, C., Mercer, N.: Children's social development, peer interaction and classroom learning. Primary Review Research Survey, 2/1b (2007)
2. Farr, W., Yuill, N., Harris, E., Hinske, S.: In my own words: configuration of tangibles, object interaction and children with autism. In: Proc. of the 9th Int. Conf. on Interaction Design and Children, pp. 30–38 (2010)
3. Lund, H.H., Jessen, C.: Playware technology for physically activating play. Artificial Life and Robotics 9(4), 165–174 (2005)
4. Shaw, M.E.: Group dynamics: The psychology of small group behavior. McGraw Hill (1971)
5. Ohyama, S., Yamada, S., Takayama, J.: Sensor node localization based on inequalities of radio field intensity - comparison of simulation and experiments. In: Proc. of SICE-ICASE International Joint Conference, pp. 1947–1952 (2006)
6. Sammon, J.W.: A nonlinear mapping for data structure analysis. IEEE Transactions on Computers C-18(5), 401–409 (1969)

Identifying Persuasive Qualities of Decentralized Peer-to-Peer Online Social Networks in Public Health

Sahiti Myneni[1,*], Sriram Iyengar[1],
Nathan K. Cobb[2], and Trevor Cohen[1]

[1] University of Texas School of Biomedical Informatics, Houston, TX, USA
{Sahiti.Myneni,M.Sriram.Iyengar,Trevor.Cohen}@uth.tmc.edu
[2] The Schroeder Institute for Tobacco Research and Policy Studies,
American Legacy Foundation, Washington, DC, USA
{NCobb}@legacyforhealth.org

Abstract. Online social networks as behavior change support systems have been rapidly gaining in popularity. These networks have been studied by public health specialists from the point of view of traditional behavior change theories. However, scant research exists on the persuasive content of the messages exchanged between participants in such networks. In this paper, we use Persuasive Systems Design (PSD) model to develop a framework for identifying persuasive attributes in online social networks. This framework was applied to QuitNet, an online social network for smoking cessation. Results indicated that the communication in QuitNet had persuasive qualities such as reduction, simulation, social learning, reminders, suggestions, and rewards. Further, these features were predominantly found in messages related to "QuitNet-specific customs", "Quit progress", and "Family and friends". Use of this framework enables the development of quantitative relationships between persuasive attributes and behavior change outcomes experienced by network members.

1 Introduction

The adverse health consequences of risky behaviors such as smoking have been well documented [1]. Several interventions have been proposed to help individuals cease smoking. Online social networks have been seen as a viable solution to promote such behavior changes [2]. These networks enable decentralized peer-peer support, which is not the case of other web-based systems. Traditionally, online social network analysis has been conducted from network science standpoint [3] or social support perspective [4]. There are few studies that analyze the persuasive attributes of communication in online social networks that promote lifestyle changes such as smoking cessation. In this paper, we analyze QuitNet, an online social network for smoking cessation using a novel framework that takes into account both behavior change theories prominent in the public health domain and persuasive systems design principles.

S. Berkovsky and J. Freyne (Eds.): PERSUASIVE 2013, LNCS 7822, pp. 155–160, 2013.

2 Theoretical Rationale

Online social networks are virtual communication platforms that allow users to share experiences, ask questions, and provide emotional support, and advice to one another. These networks form the basis for decentralized peer-to-peer support. In other words, the triggers to elicit and/or sustain a behavior change can come from multiple nodes (peers). In contrast, web-based interventions that are fully automated and technology-driven are examples of centralized support where triggers can come from only one node (the system) [5]. There can be a third kind of support infrastructure which is a hybrid of decentralized and centralized structure as found in moderated social networks, where the interactions can be peer-to-peer but there is a high-level hierarchy that guides the communication. Messages in decentralized networks do not necessarily arise from a single authority figure as is often the case with centralized systems. Thus the persuasive content arises in an organic form with many messages arising in reaction to others. Therefore, every message between decentralized network members can act as a vehicle for persuasion. In this paper, we propose a framework that enables our understanding of the persuasive qualities of a decentralized online social network. The framework involves two major steps- (a) derivation of empirically-grounded and theoretically-valid communication themes, and (b) mapping communication themes of the decentralized network to persuasive qualities.

The first step involves qualitative coding enabling us to capture the nature of communication in a decentralized network. The communication themes derived from qualitative coding are validated using existing theoretical models of behavior science and public health, thus making our analysis both empirically and theoretically-grounded. The Transtheoretical Model (TTM) [6] and Social Cognitive Theory (SCT) [7] are widely used in public health research and can be used to implement our framework. Then, the mapping process forms a bridge between theoretically-validated themes and persuasive qualities. While the first step is primarily driven by online social network content and theories of behavior change, the latter builds on persuasive system design principles. The Persuasive Systems Design (PSD) model provides a set of persuasive strategies in multiple categories [8]. In order to account for the nature of the network content, the proposed framework deals with three categories of the PSD model- primary task support, dialogue support, and social support. In this study, the framework is applied to identify the persuasive qualities embodied in communication within QuitNet, an online social network for smoking cessation.

3 Research Methods

QuitNet is one of the first online social networks for health behavior change promoting smoking cessation. Previous studies on QuitNet indicated that participation in the online community was strongly correlated with abstinence [9]. A database of 16,492 de-identified public messages between March 1, 2007 and April 30, 2007 was used in our study. All messages are stripped of identifiers but re-coded for sender id, receiver id, self-reported smoking status, and posting date.

Of the 16,492 messages, 550 messages were randomly selected for qualitative analysis. Open coding, axial coding, and constant comparison techniques [10] were used to derive themes representing QuitNet content. Examples of QuitNet themes include- "Family & Friends", "Quit obstacles", and "QuitNet-specific customs". The derived themes were validated using socio-behavioral constructs such as self-efficacy, observational learning, described in SCT and TTM. The theoretically-grounded themes were then mapped to the persuasive qualities listed under primary task support, social support, and dialogue support categories in the PSD model. It is important to note that a single message can be coded as relating to multiple themes and the persuasive qualities can overlap across themes. Table 1 shows the framework to map theoretically-grounded themes to persuasive qualities. A detailed narration of the mapping process for "Virtual rewards" theme is provided below.

The theme "Virtual rewards" consisted of messages in which members discuss virtual gifts such as bracelet to celebrate their achievements. While its connection to rewards strategy in PSD model is apparent, this theme has additional implicit persuasive attributes such as 1) it can elicit observational learning in a member when he/she watches other members receive this reward, 2) getting virtual rewards brings recognition to the network members, 3) peer recognition in turn provides a simulation cue to other members and help sustain their decision to quit, 4) it allows individuals to focus on small scale tasks, ultimately supporting them towards their target behavior, 5) the reward gets tailored and personalized based on the member's timeline. Similarly, other themes were mapped to the persuasive qualities in the PSD model. The QuitNet themes and persuasive qualities were then re-analyzed using the framework to incorporate self-reported smoking status information.

Table 1. A framework to identify persuasive qualities in a decentralized social network

QuitNet themes / Persuasive strategies	Virtual rewards	Quit obstacles	Quit benefits	Motivation	Help!	Teachable moments	Crave war	Quit progress	Social support	QuitNet-specific customs	Relapse confessions	Family & Friends
Dialogue support available												
Praise									x			x
Rewards	x											
Reminders								x		x		
Suggestion				x					x			
Similarity						x		x				
Liking												
Social role				x	x				x	x	x	x
Social Support												
Social learning	x	x	x			x	x	x		x	x	
Social comparison		x						x		x		x
Normative influence			x			x	x				x	x
Social facilitation		x			x					x		x
Cooperation					x					x		
Competition												
Recognition	x									x		
Primary Task Support												
Reduction	x							x		x		
Tunneling				x		x			x			
Tailoring	x	x	x	x	x	x	x	x	x		x	x
Personalization	x	x	x	x	x	x	x	x	x		x	x
Self-monitoring		x						x				
Simulation	x		x							x	x	
Reharsal												

4 Results

A total of 12 theoretically-grounded themes were identified in QuitNet content. These 12 themes were mapped to the persuasive qualities in the PSD model using our framework. Social learning was the most commonly found persuasive quality in QuitNet messages. On the other hand, no evidence was found for liking and rehearsal. Fig. 1(a) provides a detailed overview of the persuasive qualities in QuitNet content. A detailed distribution of the persuasive qualities in each of the QuitNet themes can be found in Fig. 1(b). "QuitNet-specific customs" themed messages had the highest number of persuasive qualities, followed by "Quit progress", "Virtual rewards", and "Family &Friends". The findings from our work concurred with the existing knowledge in behavior science research on social networks and its facilitation of social support and influence [11].

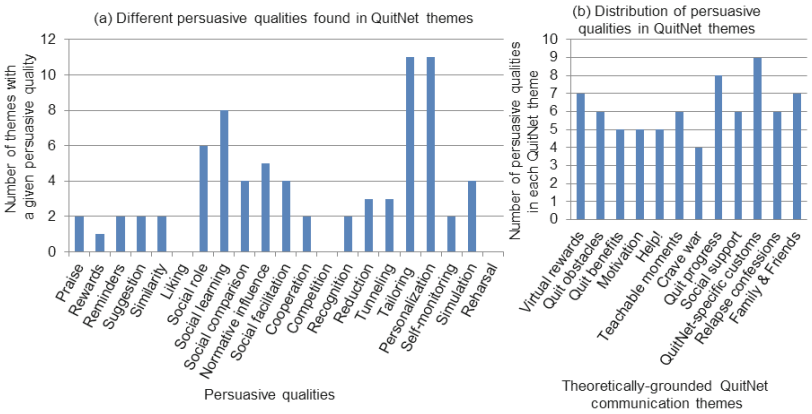


Fig. 1. Persuasive attributes and QuitNet themes

Based on self-reported smoking status, we classified QuitNet members into five groups (see Table 2). The percentage of messages sent and received by each QuitNet group across 12 QuitNet themes were calculated. Using our framework, we employed an additive model and estimated the number of persuasive qualities each group got exposed to during their QuitNet communication events as shown in Fig.2. Other models can also be easily calculated in which a persuasive quality is counted only once across themes in which it overlaps.

Table 2. Distribution of QuitNet mambers based on their self-repprted smoking status

QuitNet groups	% of messages sent	% of messages received
Smoker	20.37	25.13
Smoker to Ex-smoker	7.81	7.64
Ex-smoker to smoker	0.68	0.34
Ex-smoker	63.33	58.23
Other	7.81	8.66

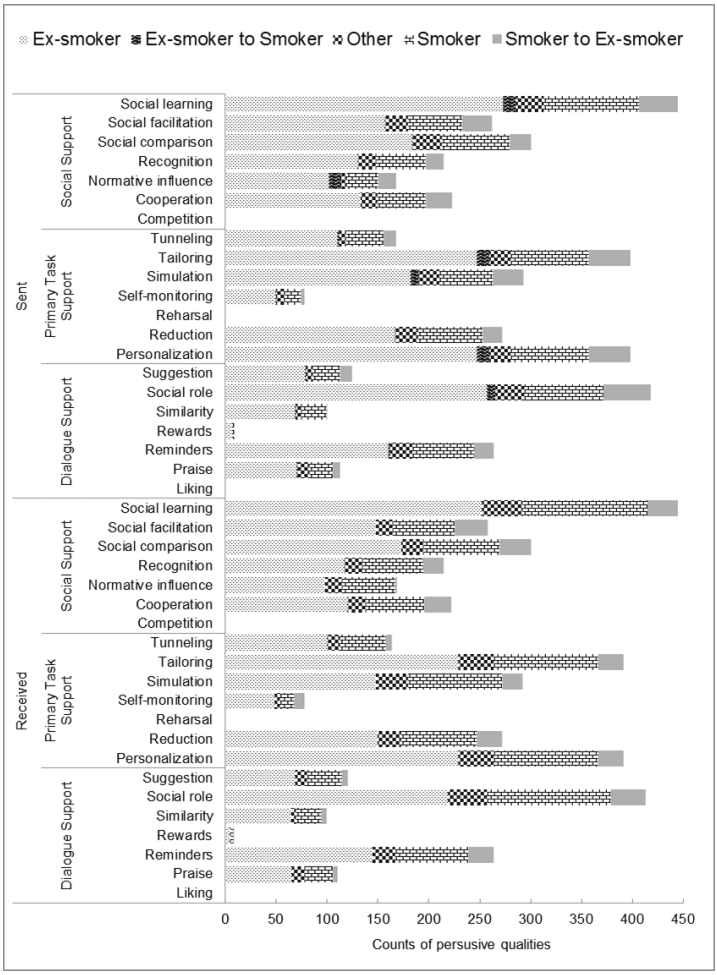


Fig. 2. Persuasive qualities and smoking status

The communication content in QuitNet possessed higher counts of persuasive qualities in the social support and primary task support categories than in the dialogue support across all messages exchanged with all QuitNet groups, except those sent from ex-smokers to smokers. The messages in the latter group had high number of persuasive qualities in the category of primary task support, followed by social support and dialogue support. Social learning, social role, personalization, and reminders were frequently found in the QuitNet messages sent and received by all five groups of QuitNet members. Given that the messages posted on decentralized networks such as QuitNet were in response to a specifc question, majority of the QuitNet communication themes were tailored and personalized, and hence these two persuasive qualities were also frequently found in QuitNet message content.

5 Conclusions and Discussion

In this paper, we analyze the persuasive qualities of communication content in QuitNet, a decentralized peer-to-peer online social network for smoking cessation. There are several benefits to the bottom-up analysis described in the paper- 1) it provides an approach to add the dimension of behavioral outcomes to the analysis of persuasive qualities, 2) it is flexible in terms of thematic granularity (theoretical constructs, persuasive categories and qualities) and analytic dimensions (communication direction, behavioral outcomes). While the results presented in this paper are preliminary and still evolving, the analysis of QuitNet demonstrates the feasibility of deducing quantitative relationships between persuasive qualities and behavior change outcomes. The proposed framework is our first step towards the analysis of persuasive qualities of the digital age communication platforms. Future studies can use this framework to assess the effectiveness of virtual support systems and translate the results into intervention development solutions.

Acknowledgement. This study is supported in part by UTHealth Innovation for Cancer Prevention Research Pre-doctoral Fellowship, UT SPH- CPRIT grant #RP101503.

Disclaimer: The content is solely the responsibility of the authors and does not necessarily represent the official views of the Cancer Prevention and Research Institute of Texas.

References

1. Organization, W.H. WHO report on the global tobacco epidemic, 2011: Warning about the dangers of tobacco. Geneva, Switzerland (2011)
2. Cobb, N.K., Graham, A.L., Byron, M.J., et al.: Online social networks and smoking cessation: a scientific research agenda. J. Med. Internet Res. 13(4) (2011)
3. Centola, D.: The spread of behavior in an online social network experiment. Science 329(5996), 1194–1197 (2010)
4. Chuang, K.Y., Yang, C.C.: A Study of Informational Support Exchanges in MedHelp Alcoholism Community. In: Yang, S.J., Greenberg, A.M., Endsley, M. (eds.) SBP 2012. LNCS, vol. 7227, pp. 9–17. Springer, Heidelberg (2012)
5. Reiter, E., Robertson, R., Osman, L.M.: Lessons from a failure: Generating tailored smoking cessation letters. Artificial Intelligence 144(1), 41–58 (2003)
6. Prochaska, J.O., Velicer, W.F.: The Transtheoretical Model of Health Behavior Change. American Journal of Health Promotion 12(1), 38–48 (1997)
7. Bandura, A.: Social cognitive theory. Encyclopedia of Psychology 7, 329–332 (2000)
8. Oinas-Kukkonen, H., Harjumaa, M.: Persuasive systems design: Key issues, process model, and system features. Communications of the AIS 24, 28 (2009)
9. Cobb, N.K., Graham, A.L., Bock, B.C., et al.: Initial evaluation of a real-world Internet smoking cessation system. Nicotine Tob. Res. 7(2), 207–216 (2005)
10. Strauss, A., Corbin, J.: Basics of Qualitative Research: Grounded Theory Procedure and Techniques. Sage, Newbury Park (1990)
11. Heaney, C.A., Israel, B.A.: Social networks and social support. Health Behavior and Health Education: Theory, Research, and Practice 3, 185–209 (2002)

Modeling Gender Differences in Healthy Eating Determinants for Persuasive Intervention Design

Rita O. Orji, Julita Vassileva, and Regan L. Mandryk

Computer Science Department
University of Saskatchewan
Saskatoon, SK. S7N 5C9, Canada
rita.orji@usask.ca, {jiv,regan}@cs.usask.ca

Abstract. The onset of many health conditions, such as obesity and type 2 diabetes, can be prevented or at least delayed by adequate changes in diet. Various determinants of healthy eating – such as *Weight Concern, Nutrition Knowledge, Concern for Disease, Social Influence,* and *Food Choice Motive* – have been manipulated by persuasive technologies to motivate healthy eating behavior. However, the relative importance and the dynamic of interaction between the determinants of healthy behavior for males and females are still unknown. Understanding how the determinants vary across user groups is important, as it will help persuasive technology designers personalize their interventions to the target demographics, thereby increasing the effectiveness of the intervention. To investigate for possible variations in healthy eating determinants, we developed separate models of healthy eating determinants for males and females. The models, which are based on a quantitative study of 228 (124 males and 104 females) participants, reveal some similarities and differences in the interactions between the determinants of healthy eating behavior. Based on the result from our models, we highlight some gender-inclusive and gender-specific approaches to persuasive intervention design.

1 Introduction

Research has shown that good eating behavior can prevent or at least reduce the risk of obesity, cardiovascular disease, certain types of diabetes, and cancer [1]. It is, therefore, not surprising that behavior change interventions aimed at modifying dietary behavior have been identified as the cornerstone treatment for these conditions [2]. Persuasive Technology (PT) is a promising approach to tackle obesity by motivating healthy behavior change using technologies. Several frameworks exist for designing and evaluating PT interventions [3]; however, these frameworks assume a one-size-fits-all approach, providing little or no information on how these strategies can be tailored to various user groups. The realization that the one-size-fits-all approach may not be sufficient for motivating healthy behavioral change has led to a growing interest in ways of tailoring interventions to various user groups and individual users. For example, previous work has shown that a user's personality is an important determinant of motivation and persuadability [4]. Further work showed a relationship between the user's personality and the success of different PT strategies [6].

S. Berkovsky and J. Freyne (Eds.): PERSUASIVE 2013, LNCS 7822, pp. 161–173, 2013.

The success of different PT strategies can be explained in part by how an individual or user group is influenced by different determinants of healthy eating. If we can understand how users are influenced by different determinants, we can effectively tailor interventions. Recent work identified *concern for weight, nutrition knowledge, concern for disease, social influence,* and *food choice motive* as determinants of healthy eating [7]; however, how the influence of these determinants varies across user groups is still unknown. For example, researchers have identified gender differences in dietary behavior [8]; however, whether or not these gender differences in eating behavior vary differentially with the determinants of healthy eating remains unclear. Understanding how the determinants vary across gender groups is important; it will help PT designers personalize their interventions to target demographics, thereby increasing the effectiveness of the PT intervention. To investigate this, we developed separate models of healthy eating determinants for males and females. The models, which are based on a quantitative study of 228 (124 males and 104 female) participants, show some similarities and differences in the relationships between the determinants of healthy eating behavior for males and females. These gender differences partially explain why males and females differ in their healthy eating behavior and suggest the need to make PT interventions gender-sensitive. We conclude by proposing ways of designing PT that are gender-inclusive (targeting both genders) and gender-specific (tailored for each group). We also shed light on the decision process in fast food eating behavior for both males and females and the need for persuasive approaches suitable for each decision process to motivate change.

2 Related Work

PT designers aim to manipulate certain behavior determinants to steer the user towards a target behavior. One way of increasing the effectiveness of a PT is to tailor it to individual users or certain user groups (based on determinants that are shown to be important for the individual or the group). The gender of a user is an important consideration for intervention tailoring. We present some gender considerations in technology design and determinants of healthy eating that has been used in PT design.

2.1 Motivations for Healthy Eating and Gender-Relevant Design

Persuasive technology that aims to bring about desirable change in attitude and behavior [9] can be effective if based on various determinants of healthy behavior. As a result research attention has focused on identifying determinants of healthy behavior, including both domain-specific and generic determinants. Recent work identified *Weight Concern, Concern for Disease, Nutrition Knowledge, Social Influence,* and *Food Choice Motive* as determinants of healthy eating [7]. These determinants can be classified into two groups: *intrinsic* and *extrinsic* determinants. In line with self-determination theory [10], intrinsic determinants, such as health optimization, involve rewards that are internal to the individual. *Weight concern,*

concern for disease, nutrition knowledge, and *food choice motive* can be classified as intrinsic determinants. Extrinsic determinants, such as approval or disapproval of others, involve external rewards or punishments [11]. *Social influence* belongs to this category. Generally speaking, each determinant might have both intrinsic and extrinsic attributes. This classification is therefore based on dominant attribute. Previous research has shown that both intrinsic and extrinsic determinants impact healthy eating behavior [7]. Consequently, various persuasive technology applications have been informed by these determinants. *Escape from Diab* [12] is a persuasive application that uses a game-based approach to motivate healthy eating and exercise. *Escape from Diab* manipulates the *weight concern* and *nutrition knowledge* to bring about desirable change. Similarly, *LunchTime* manipulates the user's *nutrition knowledge* in order to motivate healthy eating [13]. The impact of various healthy eating determinants in motivating healthy eating has been established; however, whether these determinants impact on differently males and females is unknown.

Recently, technology developers have begun to consider gender differences in their interface design decisions. For example, Moss et al. [14] reveal differences on how males and females perceive and rate websites. Their studies show that people tend to rate websites higher if designed by people of the same gender. By default (without any gender-inclusive or gender-specific design guide), male developers tend to design for male aesthetic and female developers for female aesthetics [14]. Persuasive research has also shown some gender differences in perceived credibility [15] and the moderating effect of gender on perceived persuasiveness [16]. Therefore, PT for motivating behavior change can appeal to target gender groups by employing gender specific design strategies. Therefore, examining the influence of various determinants on healthy behavior for males and females will be useful in guiding PT designers on the important determinants to focus when designing gender-tailored or gender-inclusive PT intervention.

3 Study Design and Methods

The data reported in this paper is part of a project aimed at designing technological interventions for promoting healthy eating.

We employed both quantitative and qualitative methods to examine factors affecting participants' healthy eating behavior as applied to the decision to eat fast food. We elicited participants' responses to surveys that were used to assign weightings to the five determinants of healthy behavior – *weight concern, concern for disease, nutrition knowledge, social influence,* and *food choice motive*. We were specifically interested in the relationship between the five healthy eating determinants and *health behavior* as they inform decisions making. We supplemented the survey with a 5-minute interview with 15 participants aged 18 to 36 (7 females and 8 males) aimed to elucidate the reasons behind the behaviors, and to clarify responses from the survey. In this section, we first describe how we developed the research instrument; this is followed by a description of the data collection methods used and the validation of our analysis.

3.1 Measurement Instrument

To collect data for our model, we developed a survey to collect responses from customers from a set of fast food restaurants in Canada. "Fast food" is defined here as convenience foods obtained in self-service or 'take away' restaurants / eateries with minimal waiting. These foods are characterized with high energy density, low levels of micronutrients and fiber, high levels of simple sugars and salt; they are usually sold in larger portion sizes than conventional home-cooked or restaurant foods [17]. The survey was developed after a detailed review and pilot tested (n=10) for refinement.

The survey instrument consisted of questions assessing participants' (1) demographics (gender, age group, education); (2) health concern (weight concern and concern for diseases); (3) nutrition knowledge; (4) fast food motives; (5) healthy eating attitude; (6) frequency of purchase; (7) purchase decision type, and (8) social influence. We adapted the 10 questions from Kahkonen [18] to measure the health concern variable. The scale has been validated by several studies [18, 19]. The health related concern measures the participants' degree of concern about food and health related issues using 5-point Likert scale 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: Weight Concern (WC) and Concern for Disease (DC). This is based on the factor loadings and suggestions by previous studies [7, 19], that WC and DC might impact 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. [20]. Some examples include "It is convenient", "It is healthy", and "It is cheap". Our survey allowed the participants to select from a list of factors that motivate their fast food consumption, which were mostly non-health related. We measured attitude towards healthy eating using a 3-item scale adapted from Kearney et al. [21]. An example of a question in this category is "I make a 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 [22] 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?" The social influence variable was included 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. Also, the

purchase decision question was added to examine if decision to eat fast food is an impulsive or a planned decision. This is important because decision on suitable persuasive approach might differ depending on whether target behavior involves an impulsive or a planned decision.

3.2 Participants

The participants consisted of 228 restaurant visitors sampled at selected fast food restaurants within a Canadian University. The data were gathered over a period of 30 days. 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). There was a relatively even distribution of females (46%: 104 participants) and males (54%: 124 participants). The ages of the participants were unevenly distributed: 18-25 (155, 68%), 26-35 (59, 26%), 36-45 (11, 5%), and only 3 (1%) over 45. Similarly, most of the participants in this study were high school graduates who are presently pursuing their first degrees 104 (46%); 5 (2%) held a diploma degree, 53 (23%) held a bachelor's degree, 57 (25%) were master's degree holders, 7 (3%) were doctorate degree holders and, 2 (1%) held other forms of certificates. Regarding the frequency of visits to fast food restaurants, more than 60% of the participants visited fast food restaurants at least 2 times a week (every day: 8, 4%; 3 times a week: 62, 27%; once a week: 73, 32%; 3 times a month: 32, 14%; once a month: 39, 17%; 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. A significant number of females (66: 63%) make an impulsive decision to buy fast food, while 85 (69%) males indicated that their fast food consumption is often a planned decision.

3.3 Measurement Validation

To determine the validity of our survey instrument we conducted an Exploratory Factor Analysis (EFA) in SPSS 19 using Principal Component Analysis (PCA).

Confirmatory Factor Analysis

We identified five determinants of healthy eating from the literature – *weight concern, disease concern, nutrition knowledge, social influence,* and *food choice motive*. To verify that our data replicated the five factors in healthy eating behavior, we conducted Confirmatory Factor Analysis (CFA – a statistical procedure that compares the fit of the data with the factor being modeled) using Partial Least Square (PLS) Structural Equation Modeling (SEM). PLS is especially recommended for theory formation and verification [24]. Moreover, PLS-SEM has less stringent requirements concerning data distribution assumptions [24]. In the CFA, the five factors were included as latent (independent) variables, and each was hypothesized to have a direct effect on healthy eating attitude – the dependent variable. We report the

recommended set of indices for model validity and reliability in PLS. We used SmartPLS 2 [25] to analyze the model.

Indicator reliability can be assumed because Cronbach's α and the composite reliability that analyze the strength of each indicator's correlation with their variables are all higher than their threshold value of 0.7 and 0.6 respectively [26]. Convergent and discriminate validity can be assumed as all constructs have an Average Variance Extracted (AVE) – which represents the variance extracted by the variables from its indicator items – above the recommended threshold of 0.5 and greater than the variance shared with other variables [26]. The models yield acceptable values of all indices for PLS model validity or reliability. Again, prior to comparing our two models, we established measurement invariance between males and females samples. The psychometric properties from our two groups demonstrate that the two groups have the same structure; therefore our data is suitable to conduct group comparison. Measurement Invariance was assessed using component-based CFA via SmartPLS [25].

4 Results and Discussion

To examine the differences and similarities between males and females (in terms of their healthy eating determinants), we systematically examined the interactions and the influences of the five determinants (*weight concern, concern for disease, nutrition knowledge,* and *social influence*) on healthy eating behavior using SEM in SmartPLS.

4.1 The Structural Model

The structural model determines the interactions between determinants – specifically, the relationship between the determinants and healthy eating attitude. Figures 1 and 2 show the structural models of healthy eating determinants for males and females respectively. An important criterion to measure the strength of the relationship between variables in structural models is the level of the path coefficient (β) and the significance of the path coefficient (p) [27]. Path coefficients measure the influence of one variable on another. The individual path coefficients and their corresponding level of significance obtained from the two models are shown in Figures 1 and 2.

4.2 Comparison of Eating Determinants for Females and Males

The results from our models show that males and females differ with regards to the influence of the determinants (*weight concern, concern for disease, nutrition knowledge,* and *social influence*) on their healthy eating behavior (see Figures 1 and 2). Females show a more holistic approach to healthy eating behavior in relation to the influence of the determinants. A pairwise comparison of multi-group differences reveals some significant differences ($p<0.05$) in the interaction between the determinants and healthy eating attitude in the two groups. We discuss and compare the influence of the determinants on males and females.

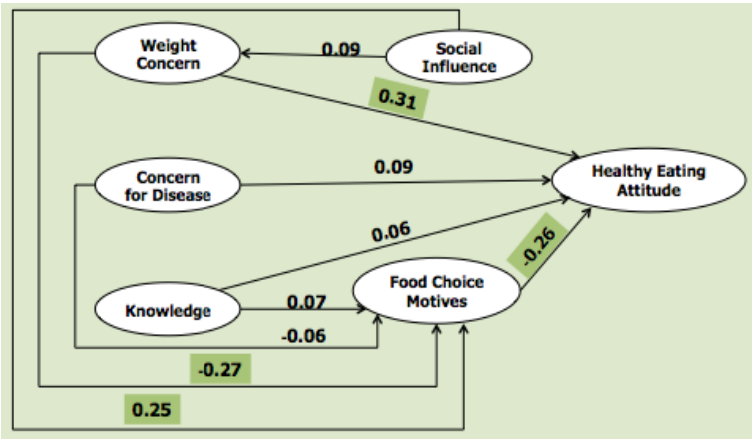


Fig. 1. A model of healthy eating attitude and associated determinants' interactions for males. Significant coefficients are highlighted

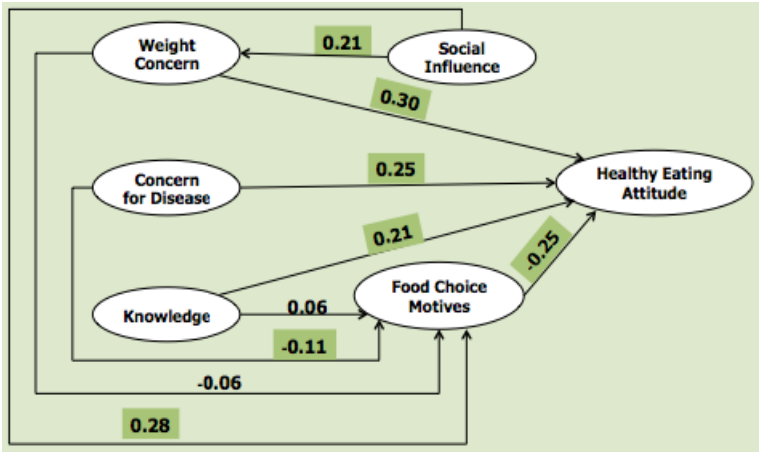


Fig. 2. A model of healthy eating attitude and associated determinants' interactions for females. Significant coefficients are highlighted

Weight Concern: People may adopt healthy behaviors for reasons that are unrelated to health [7]. Being overweight has many negative health-related implications; however, the post survey interviews revealed that the participants were mostly concerned about their weight because they want to maintain an attractive appearance. Our model shows that *weight concern* is a significant positive determinant of healthy eating behavior for both males and females ($\beta = .31$ and $\beta = .30$, respectively, $p<.01$) as shown in Figures 1 and 2. This is in line with the previous finding that *weight concern* is an important consideration in people's decision to adopt a healthy eating behavior [7]. A possible explanation of this finding is the fact that calories (a major source of weight) can only be obtained by consuming food. Therefore, people who are concerned about their weight can easily be persuaded to adopt a healthy eating

behavior to control their weight. A female participant in our post survey related how important it was for her to maintain an 'ideal' weight because it served as a basis for her self-esteem.

Also noteworthy is the fact that, although, *weight concern* affected healthy eating attitude positively for both males and females, its effect on *food choice motive* is only significant for males (β = -.27, p<.01). This suggests that males that are concerned about gaining weight from consuming unhealthy food not only adopt a healthier eating attitude, but also show more resistance to the barriers on healthy eating imposed by *food choice motive*. Therefore, an effective way of overcoming the inhibiting effect of barrier to healthy eating posed by *food choice motive* is to design interventions to increase the perceived relationship between unhealthy eating and weight gain especially for males. Overall, the relationship between *weight concern* and healthy eating attitude implies that persuasive intervention designers aiming at developing gender-inclusive applications for motivating healthy eating attitude can emphasize and make obvious the relationship between unhealthy eating and weight gain to motivate both genders. For instance, a tracking and monitoring application can portray the outcome of healthy behavior in terms of the amount of weight lost (or gained in case of unhealthy behavior) to motivate both males and females to adopt healthy behavior.

Concern for Disease: Previous research on the influence of *concern for disease* on healthy eating attitude has generated contradicting views and results. Sun [19] showed that most individuals expressed concern about unhealthy eating and developing food-related chronic diseases while Orji et al. [7] shows that *concern for disease* is not a significant determinant of a healthy eating attitude. These mixed findings are not surprising considering that none of the studies have considered other moderating factors that might alter the relationship between *concern for disease* and eating attitude. Our models show that the influence of *concern for disease* on attitude is in fact influenced differently by gender. *Concern for disease* is a significant motivator of healthy eating attitude for females (β = .25, p<.01) but not for males as shown in Figures 1 and 2. A possible explanation why *concern for disease* is not as strong a motivator as *weight concern* for both males and females is that there are so many other factors that can cause diseases other than unhealthy eating. Similarly, research has shown that people differ in their beliefs about diet-disease relationship and attributions of illness (its causes, cures, treatments, and symptoms) [28]. Our results suggest that males do not consider the relationship between one's dietary behavior and diseases, while females do. This is in line with a previous study that shows that Canadian male students engage in more risky health-related behaviors to show strength, fearlessness, and manliness [8]. Men perceive their health status as excellent because they tend to consider themselves invulnerable to a number of potential health threats [8]. This implies that persuasive interventions attempting to motivate healthy behavior by manipulating the relationship between (un)healthy eating and diseases might not be effective for males, whereas they are likely to motivate a positive behavior change in females. The mediating role played by *food choice motive* on *concern for disease* for females is worth noting. The negative influence of disease concern on *food choice motive* (β = -.11, p<.01) implies that designing persuasive technology to increase an individual's consciousness of acquiring diseases from

unhealthy eating is a promising way of reducing the inhibiting effect of *food choice motive* on healthy eating attitude for females.

Nutrition Knowledge: Previous research has shown that knowledge influences healthy eating attitude positively [7]. Similarly, research examining the relationship between *nutrition knowledge* and gender shows that females possess a higher level of *nutrition knowledge* than males [29]. Consequently, several PT interventions have been designed based on the idea that providing *nutrition knowledge* will go a long way towards promoting healthy eating behavior [13]. Our model, however, reveals that knowledge is only a significant determinant of healthy eating attitude for females ($\beta = .21$, $p<.01$). In fact designing persuasive interventions to impact *nutrition knowledge* on males might have no direct or indirect influence on their healthy eating attitude. This supports previous research that shows that knowledge alone might not be sufficient to cause a desired behavior change [30]. Persuasive interventions targeting males might not focus on any strategy that impact *nutrition knowledge* as a key motivator of change.

Social Influence: In line with previous research [7], *social influence* has no direct influence on healthy eating attitude for both males and females. However it strengthens the influence of *food choice motive* for both males and females ($\beta = .25$ and $\beta = .28$, respectively, both $p<.01$). This implies that the hindering effect of *food choice motive* on healthy attitude is amplified by related others (e.g., peers). Also noteworthy is the fact that *social influence* increases the effectiveness of *weight concern* for females. This is in line with previous findings that social support provided by friends and family members can enhance both weight loss and maintenance [31]. This result can be explained by the fact that friends eat together as a way to form social bonding. As noted in previous work [7], fast food restaurants are convenient, fast, and inexpensive places for group eating. Our model results also suggest that *social influence* can facilitate or hinder healthy behavior. *Social influence* can hinder the tendency of healthy eating by strengthening the negative influence of *food choice motive* on healthy eating attitude; however, it also amplified concern for gaining weight from unhealthy eating for females. The findings show that *social influence* may not be used as a key motivator of behavior change in a persuasive intervention but can be used to amplify the effect of other motivators like *weight concern*.

Food Choice Motive: *Food choice motive* is the only determinant that influences healthy eating attitude negatively for both males and females ($\beta = -.26$ and $\beta = -.25$, respectively, $p<.01$). This is as expected as most of the motives behind fast food consumption are non-health related (e.g., cheap, convenient, time saving). PT designers should therefore plan to deal with the inhibiting effect of *food choice motive* on healthy eating attitude in their design. From our model, one way to reduce the negative effect of *food choice motive* is to increase *weight concern* (for males) by making obvious the diet-weight relationship. For females, however, decreasing the negative effect of *food choice motive* can be achieved by increasing *concern for disease*, which can be achieved by portraying diet-disease relationships.

Impulsive and Planning Decision: Another important discovery that emerged from our study is that the decision to eat fast food is more an impulsive decision for females (63% of females make impulsive decisions to eat fast food) and a planned

decision for males (69% of males make planned decision to eat fast food). Impulsive decisions to purchase occur when an individual makes an unintended, unreflective, and immediate purchase [32]. This result is in line with a previous study that shows that men possess goal-oriented characteristics, such as decisiveness, whereas women show more emotive qualities, such as emotional responsiveness [33]. This implies that different persuasive approaches maybe employed for males and females to cause an effective change in healthy behavior especially in relation to fast food consumption. For example, persuasive approaches that appeal to an individual's emotion (rather than his/her sense of reasoning) might be more effective for women at changing behaviors associated with impulsiveness and shallow processing of information. Logical strategies (e.g., goal setting) that appeals to a person's sense of reasoning might be more effective at causing behavior change in males.

4.3 Gender-Inclusive PT Design

Our findings can be applied to design persuasive intervention targeting both males and females (gender-inclusive) without demotivating anyone. Our results show that *weight concern* influences both males and females positively. Therefore, to appeal to a broader audience, persuasive intervention designers should include persuasive strategies that manipulate *weight concern.* For example, self-monitoring technologies can be used to track a user's behavior and the outcome can be projected in terms of weight gain or loss. Similarly *concerns for disease and nutrition knowledge* are significant motivators for females, but do not influence males negatively. Including these determinants in PT intervention design will motivate a positive behavior change for females without hindering behavior performance for males. *Social Influence* has no direct influence on attitude; however, it influences *weight concern* positively for females and therefore can also be employed. *Food choice motive* impacts on attitude negatively for both males and females. To deal with this, PT designers can increase the perceived *weight concern* and *concern for disease* in a gender-inclusive design.

4.4 Gender-Specific PT Design

Although designing gender-inclusive PT intervention is possible, personalizing persuasive interventions is advocated to increase the effectiveness of interventions and to keep the intervention simple – focusing on only determinants that are of significant to each group.

To design persuasive interventions targeting males, PT designers should employ any persuasive strategies that emphasize *weight concern* to directly impact positive attitude change. *Concern for disease* could also be employed to reduce the inhibiting effect of *food choice motive* on attitude. On the other hand, to design for females, PT designers should employ strategies that emphasize *weight concern, concern for disease,* and *nutrition knowledge* to motivate a positive change in attitude. *Social influence* can be employed to strengthen the impact of *weight concern* on attitude. Similarly, *concern for disease* can be employed to reduce the negative effect of *food choice motive* on healthy eating attitude.

Our findings can be applied in designing both gender-inclusive and gender-specific PT interventions that manipulate only the determinants that are important to the users.

Limitations and Future Work: 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. In the future, we hope to apply the findings from this study in designing both gender-inclusive and gender-specific PT interventions for motivating healthy eating behavior.

5 Conclusion

Several research findings have identified gender differences in healthy behavior especially healthy eating behavior; however, whether there are also differences in the underlying determinants of healthy eating behavior was not clear. Our study shows the existence of gender differences in the underlying determinants of healthy eating and their interactions. This finding established the need to tailor PT intervention based on gender – by targeting determinants that are significantly important to males and females. Our models show that females have a more realistic view of health behavior (in relation to its determinants). Finally, based on the results from our study, we highlighted approaches for designing both gender-inclusive and gender-specific interventions. Our study also shows that a rational persuasive approach may be more effective for males while emotional appeal may work better for females. We shed light on ways of designing gender-sensitive PT for increased effectiveness.

Acknowledgements. The first author of this paper is being sponsored by NSERC 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 (2006)
2. Lau, et al.: 2006 Canadian clinical practice guidelines on the management and prevention of obesity in adults and children. CMAJ 176, S1–S13 (2007)
3. 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)
4. 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)
5. Hu, R., Pu, P.: Acceptance Issues of Personality-based Recommender Systems. In: Proc. ACM Conf. on Recommender Systems, pp. 221–224 (2009)
6. Halko, S., Kientz, J.A.: Personality and Persuasive Technology: An Exploratory Study on Health-Promoting Mobile Applications. In: Ploug, T., Hasle, P., Oinas-Kukkonen, H. (eds.) PERSUASIVE 2010. LNCS, vol. 6137, pp. 150–161. Springer, Heidelberg (2010)

7. Orji, R., Mandryk, R.L., Vassileva, J.: Towards a Data-Driven Approach to Intervention Design: A Predictive Path Model of Healthy Eating Determinants. In: Bang, M., Ragnemalm, E.L. (eds.) PERSUASIVE 2012. LNCS, vol. 7284, pp. 203–214. Springer, Heidelberg (2012)
8. Dawson, et al.: Examining gender differences in the health behaviors of Canadian university students. J. of the Royal Society for the Promotion of Health 127, 38–44 (2007)
9. Fogg, B.J.: Persuasive Technology: Using Computers to Change What We Think and Do. M.K. Publishing (2003)
10. Deci, E.L., Ryan, R.M.: Intrinsic Motivation and Self-Determination in Human Behavior (Perspectives in Social Psychology). Plenum Press (1985)
11. Michaelidou, N., Christodoulides, G., Torova, K.: Determinants of healthy eating: a cross-national study on motives and barriers. J. of Consumer Studies 36, 17–22 (2012)
12. Thompson, D., et al.: Serious Video Games for Health How Behavioral Science Guided the Development of a Serious Video Game. Simulation & Gaming 41, 587–606 (2010)
13. Orji, R., Vassileva, J., Mandryk, R.L.: LunchTime: a slow-casual game for long-term dietary behavior change. J. of Personal and Ubiquitous Computing (2012), doi:10.1007/s00779-012-0590-6
14. Moss, et al.: Some men like it black, some women like it pink: consumer implications of differences in male and female website design. J. of Consumer Beh. 5, 328–341 (2006)
15. Ferebee, S.: The Influence of Gender and Involvement Level on the Perceived Credibility of Web Sites. In: Oinas-Kukkonen, H., Hasle, P., Harjumaa, M., Segerståhl, K., Øhrstrøm, P. (eds.) PERSUASIVE 2008. LNCS, vol. 5033, pp. 279–282. Springer, Heidelberg (2008)
16. Drozd, F., Lehto, T., Oinas-Kukkonen, H.: Exploring Perceived Persuasiveness of a Behavior Change Support System: A Structural Model. In: Bang, M., Ragnemalm, E.L. (eds.) PERSUASIVE 2012. LNCS, vol. 7284, pp. 157–168. Springer, Heidelberg (2012)
17. Feeley, A.B.B., Pettifor, J.M., Norris, S.: Fast food consumption among 17 year olds in the Birth to Twenty Cohort. South African J. of Clin Nutri. 22(3) (2009)
18. Steimann, F., Kahkonen, P., Tuorila, H., Rita, H.: How information enhances acceptability of a low-fat spread. J. Food Quality and Preference 7, 8 (1996)
19. Sun, Y.-H.C.: Health concern, food choice motives, and attitudes toward healthy eating: the mediating role of food choice motives. Appetite 51, 42–49 (2008)
20. Steptoe, A., Pollard, T.M., Wardle, J.: Development of a measure of the motives underlying the selection of food: the food choice questionnaire. Appetite 25, 267–284 (1995)
21. 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 4, 1117–1126 (2001)
22. 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)
23. Hinton, R.H., Brownlow, C., McMurray, I., Cozens, B.: SPSS Explained. Routledge (2004)
24. Henseler, J., Ringle, C.M., Sinkovics, R.R.: The use of partial least squares path modeling in international marketing. J. Advances in Internl. Marketing 20, 277–319 (2009)
25. Ringle, C.M., Wende, S., Becker, J.: smartpls.de - next generation path modeling, `http://www.smartpls.de/forum/contact.php.hp` (accessed August 2012)
26. Chin, W.W.: The Partial Least Squares Approach to Structural Equation Modeling. Modern Methods for Business Research, 298–336 (1998)
27. Hair, et al.: PLS-SEM:Indeed a Silver Bullet. J. Markting Theory and Pract. 19,139–152 (2011)

28. Vaughn, et al.: Cultural Health Attributions, Beliefs, and Practices: Effects on Healthcare and Medical Education. J. Open Medical Education J. 2, 64–74 (2009)
29. Whitcombe, et al.: Relationship between Nutrition Knowledge and Gender, Ethnic Background, and Year in School in varsity athletes. J. Amer. Dietetic Ass. 96, A87 (1996)
30. Brown, K., McIlveen, H., Strugnell, C.: Nutritional awareness and food preferences of young consumers. J. Nutrition & Food Science 30, 230–235 (2000)
31. Wing, R.R., Jeffery, R.W.: Benefits of recruiting participants with friends and increasing social support for weight loss and maintenance. J. Consult. and Clini. Psy. 67, 132–138 (1999)
32. Rana, S., Tirthani, J.: Effect of Education, Income and Gender on Impulsive Buying Among Indian Consumer an Empirical Study of Readymade Garment Customers. J. Management 145–146 (2012)
33. Rosenkrantz, P., Bee, H., Vogel, S., Broverman, I.: Sex-role stereotypes and self-concepts in college students. J. of Consulting and Clinical Psych. 32, 287–295 (1968)

Social Stairs: Taking the Piano Staircase towards Long-Term Behavioral Change

Michel Peeters, Carl Megens, Elise van den Hoven,
Caroline Hummels, and Aarnout Brombacher

Faculty of Industrial Design, Eindhoven University of Technology
Den Dolech 2, 5612AZ Eindhoven, The Netherlands
{m.m.r.peeters,c.j.p.g.megens,e.v.d.hoven,c.c.m.hummels,
a.c.brombacher}@tue.nl

Abstract. This paper addresses the development of Social Stairs, an intelligent musical staircase to change people's behavior in the long-term to take the stairs in favor of the elevator. Through designing with the Experiential Design Landscape (EDL) method, a design opportunity was found that social engagement encouraged people to take the stairs at work in favor of the elevator. To encourage this social behavior, people who involved each other and worked together whilst using the Social Stairs were treated with more diverse orchestral chimes that echoed up the stairwell. In this paper we reflect on the differences between the persuasive system of the well-known Piano Staircase and the Social Stairs. We report on the deployment of the Social Stairs for a period of three weeks in the public space within the university community and identify opportunities for triggering intrinsic motivation, social engagement and how to keep people involved in the long-term.

1 Introduction

Modern societal trends and technologies enable us to live our lives with comfort and fun. While these developments create quality of life, with respect to our health and wellbeing there are downsides as well. In our current society we are dealing with an increasing sedentary and inactive lifestyle with all the consequences (e.g. obesity, diabetes) that entail for our health and wellbeing (Schroeder, 2007).

Persuasive technology is often used and applied to provide solutions, to persuade people to become more active again. One of the pinnacle examples of persuasive technology is the Piano Staircase (Volkswagen, 2009)(see figure 1), a design concept of a musical staircase, which was built next to the escalator at the Odenplan subway station in Stockholm, Sweden. People could choose to either use the stairs or the escalator, with the latter being the more popular but unhealthier choice.

The designers of the Piano Staircase decided to try to change people's behavior by transforming the stairs into a giant piano keyboard. By applying pressure on each step the staircase would play a musical note. The resulting video of the Piano Staircase shows commuters discovering the Piano Staircase, getting curious and triggered towards the behavior of staircase by playing with the steps and eventually using the

S. Berkovsky and J. Freyne (Eds.): PERSUASIVE 2013, LNCS 7822, pp. 174–179, 2013.

stairs more often than the escalator. Through the interactive steps the staircase successfully persuaded and motivated 66% more people to use the stairs instead of the escalator throughout the day they shot the video (Volkswagen, 2009).

The Piano Staircase is persuasive technology designed with an explicit intention to change people's attitude or behavior to choose the stairs over the escalator (IJsselsteijn et al., 2006)(Lockton et al. 2010). It clearly shows the potential of how technology can persuade people, by triggering new motivations, to change an attitude, belief or behavior through persuasion and social influence, not through coercion (Fogg 2002).

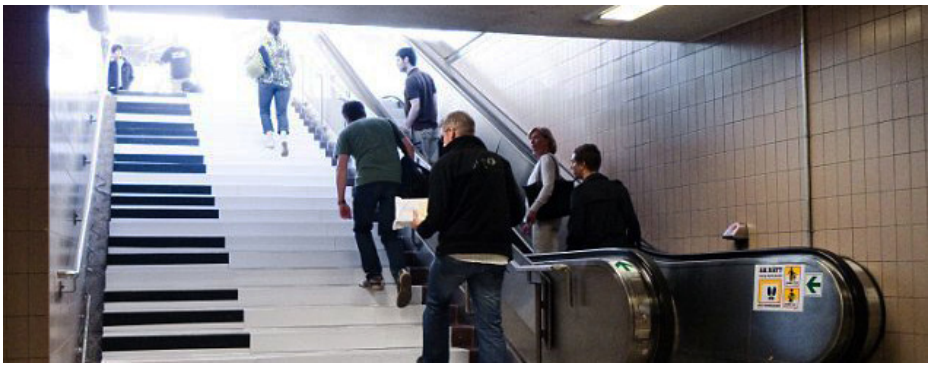


Fig. 1. Piano Staircase; applying pressure on each step plays a musical note

For the case of designing for wellbeing towards a more active and healthy lifestyle we fully support the attempts and believes of the Persuasive Technology field that behavioral change is favorable for a healthy society. We question however whether the current examples like the Piano Staircase will lead to long-term structural behavioral change. The Piano Staircase really made a difference on the first day, the moment people got introduced to this new concept and it raised their curiosity and motivation. What is not shown in the video is what happens when the same person passes the stairs every day. Will he/she still be triggered by the Piano Staircase in the long run to play with it and actually use the stairs instead of the escalator? Could this persuasion lead to a long-term intrinsic motivation for people to keep using the stairs?

2 Designing for Behavioral Change in the Long-Term

BJ Fogg's (2009) behavior model describes three elements that must converge at the same moment for a behavior to occur: Motivation, Ability, and Trigger. When behavior change does not occur, at least one of those three elements is missing. IJsselsteijn et al. (2006) point out that changing human behavior typically takes a long time and that there are, to their knowledge, "hardly any user studies available that have looked at the long-term effectiveness of persuasive technology". To explore long-term behavioral change through design we did a longitudinal experiment. For this we developed and used an Experiential Design Landscape (EDL) (Gent et al.

2011), an infrastructure where all stakeholders work together, creating experienceable propositions together with them, which evolve over time.

These propositions, Experiential Probes (EPs), are intelligent, open, sensor-enhanced, networked products-service systems that enable people to develop new and emerging behavior and, in parallel, enable detailed analysis of the emerging data patterns by researchers and designers as a source of inspiration for the development of future systems, products and services. We were interested in how design can evoke emerging or changing behavior that will lead to people's intrinsic motivation in the long-term to take the stairs in favor of the elevator. For this purpose we created an EDL called Social Stairs, where we developed and evaluated several Experiential Probes over time.

2.1 Social Stairs

In the university's main building, sensors and sound actuators were installed to the main staircase as well as a camera to observe the EDL (see figure 2). The overall goal was to get people to walk past the elevator and use the stairs instead. At first ascending and descending sounds were mapped to the steps, similar to the Piano Staircase, but not directly mimicking a piano. This, similar as on the Piano Staircase, triggered people to play with the steps, walking over the stairs in different speeds and just randomly stepping and jumping on them. The sound echoed past the elevator and raised awareness, triggering people to take a look. In contradiction to the Piano Staircase the choice between the elevator and stairs is not presented equally, as the stairs are located behind the elevator.

Fig. 2. People play together in the Social Stairs EDL

In early experiments it was found that social engagement encouraged more active behavior. One of the noticeable occurrences was that a lot of people played together with the staircase, trying to create sounds or even music together. After a few days the designers saw that people were inviting peers to join them on the stairs, helping to promote the staircase to a larger audience.

As the design is not final in an EDL, it is treated as a probe, several alterations or complete new attempts can be made by the designers to steer or find the behavior they envisioned. In this case the designers wanted to encourage this social behavior. People who worked together whilst using the Social Stairs were treated with a richer louder, orchestral chime that echoed up the stairwell. This resulted in a richer dynamic experience; the stairs sounded different depending on the situation the people were in.

2.2 Behavior on Social Stairs

Social Stairs provided the designers with user data of three weeks (i.e. log data of steps, interviews and video), which were utilized, to continuously do design iterations but also to analyze and test whether the intended effects were actually met. Through the data (see figure 3) they got insight in different types of behavior. For instance, people invited others to join them at the Social Stairs and create a soundscape together. Other people were actively seeking opportunities to create a joined soundscape, by patiently waiting for a while in the stairwell.

Unexpected and emergent behavior also occurred; some people were meeting up in the stairwell on a daily basis, similar to a hangout, and formed medium to large groups (2-10 people) to create a gigantic joined soundscape. Others got to know new people through the Social Stairs.

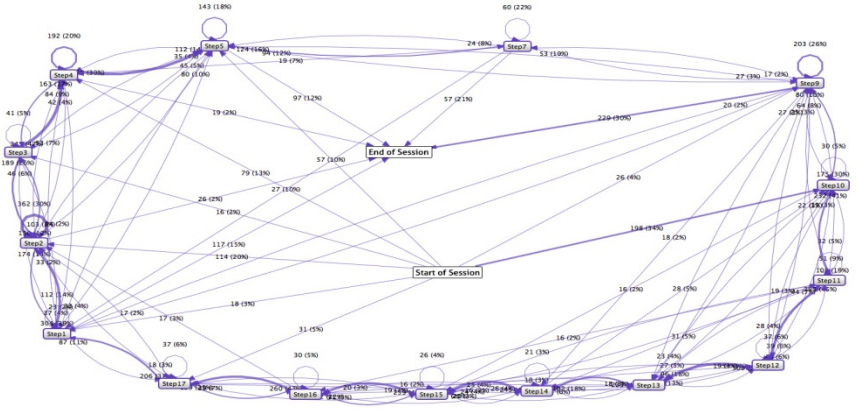


Fig. 3. Through these visualizations insight can be gained in the amount of playfulness that occurred that day. Low sequential steps and more loops means people played around and jumped around in a non-sequential order. The visualization is made in such a way that it for instance becomes easy to compare different days to each other, showing the triggered playfulness of different Experiential Probes.

What the designers found particularly interesting was the example of some recurring visitors who were driven by curiosity about possible new types of behavioral layers of the Social Stairs. They explained to the designers that they actually hadn't made a deliberate decision to take the stairs over the elevator but they

were just curious if the Social Stairs was still working and if there was something going on in the stairwell. Initially they were driven by curiosity but taking the stairs also became a new habit for them. It's even questionable if they were still deciding between the stairs and the elevator, by choosing the stairs in favor of the elevator.

3 Probing towards Intrinsic Motivation to Take the Stairs

When we look at BJ Fogg's (2009) behavior Model, the Social Stairs changed very little to people's 'ability' to physically climb the stairs or their ability to choose the stairs in favor of the elevator. The two other important elements for a behavior to occur i.e. 'trigger' and 'motivation' however did change (still with the understanding that both elements have to take place at the same moment together with people's ability, for an eventual behavior to occur). Social Stairs was a result of an EDL in which the triggers were 'open scripts' which could evolve over time; they were adapted or sometimes even completely (re)designed based on the logged data and video of the Social Stairs. With the Social Stairs we wanted to design and probe towards triggering people's intrinsic motivation to take the stairs.

With intrinsic motivation we would like to refer to the work of White (1959) and Deci (1975) where intrinsic motivation refers to motivation to enact a behavior for its inherent satisfaction, in alignment with one's personal values or attitudes, and not for a separable external consequence. Moreover, intrinsic motivation seems to increase the likelihood of the behavior being performed (Deci, 1975) and seems to lead to sustained behavior (Deci, 1975). With EDLs we therefore also aim to design towards sustainable behavioral change through triggering people's intrinsic motivation. As people may lose attention to the act of persuasion we therefore believe that true sustainable behavioral change most likely has to come from within people themselves to last.

Therefore, designers should focus their designs to help people trigger intrinsic motivation towards changing behavior, not persuade and extrinsically motivate them to act or behave different in that moment of time. It is however very hard to predict which trigger will lead to intrinsic motivation on the long term. For designers it is almost impossible to predict what behavior will emerge or change once people start to use their designs. By using the EDL method designers can probe and find out during the process of design.

4 Discussion

The EDL method allows designers to explore and play around with their designs and people's resulting behavior. At this moment the Social Stairs EDL is still ongoing and much is left to explore. The designers are further developing and exploring ways to make the sounds more dynamic. Several probes are created in which people over time can change the sounds, both the character of the sound itself, but also through repeating loops of recent visitors creating an echo-effect. Aim is to create a multi-layered experience, in which people can continuously (and jointly) explore and

deepen the interaction with the Social Stairs. This way the designers aim to design for long-term structural new behavior change instead of 'just' renewing the sounds of the Social Stairs.

Next to this, research has to be done on the data outcome of the stairs. In the first three weeks of the EDL a lot of data has been accumulated which needs further and deeper analysis. Already an analysis of playful behavior has been done (see figure 3), but we believe there is more understanding of behavior to be gained from the combination of the log data, video and interviews.

Nevertheless the EDL method shows to be promising for long term, structural behavior change. Through the combination of continuously analyzing and designing in the EDL designers can react and anticipate on (unexpected) behavior of the people involved.

Acknowledgments. Ministry of Economic Affairs under the IOP-IPCR program. We would like to thank the students involved in this project, responsible for the design of the probes: Nadine van Amersvoort, Rhys Duindam, Nick Hermans, Max Sakovich & Bart Wolfs.

References

1. Deci, E.L.: Intrinsic motivation. Plenum, New York (1975)
2. Fogg, B.J.: Persuasive technologies: Introduction. Communications of the ACM 42 (1999)
3. Fogg, B.J.: Persuasive Technology: Using Computers to Change What We Think and Do. Morgan Kaufmann Publishers, San Francisco (2002)
4. Fogg, B.J.: A behavior model for persuasive design. In: Proceedings of the 4th International Conference on Persuasive Technology
5. van Gent, S.H., Megens, C.J.P.G., Peeters, M.M.R., Hummels, C.C.M., Lu, Y., Brombacher, A.C.: Experiential design landscapes as a design tool for market research of disruptive intelligent systems. In: Proceedings of the 1st Cambridge Academic Design Management Conference. University of Cambridge, Cambridge (2011)
6. IJsselsteijn, W.A., de Kort, Y.A.W., Midden, C., Eggen, B., van den Hoven, E. (eds.): PERSUASIVE 2006. LNCS, vol. 3962. Springer, Heidelberg (2006)
7. Lockton, D., Harrison, D., Stanton, N.: The Design with Intent Method: A design tool for influencing user behavior. Applied Ergonomics 41(3), 382–392 (2010)
8. Schroeder, S.: We Can Do Better - Improving the health of the American people. The New England Journal of Medicine 357, 1221–1228 (2007)
9. Volkswagen Group Sverige AB 2009, Sodertalje, `http://www.thefuntheory.com/piano-staircase` (accessed January 21, 2013)
10. White, R.W.: Motivation reconsidered. Psychol. Rev. 66, 297–333 (1959)

Invisible Work: An Ambient System for Awareness and Reflection of Household Tasks

Wolfgang Reitberger, Martin Kastenmiller, and Geraldine Fitzpatrick

Institute for Design & Assessment of Technology, Vienna University of Technology, Argentinierstraße 8, 1040 Vienna, Austria
{wolfgang.reitberger,geraldine.fitzpatrick}@tuwien.ac.at, kastenmiller@gmail.com

Abstract. Household tasks have been described as the invisible work carried out in the home. Their coordination and especially negotiating a fair and transparent distribution of these tasks between different members of a household is a significant challenge in the busy lives of many people. The persuasive system presented in this paper, Choreflect, aims to address this issue by making the inhabitants' contributions to household work more visible. The prototype uses distributed ambient displays in the home to enable household members to track household tasks, which are then visualized by the system. A four week study of the system carried out in two shared apartments and two family households showed that the system fostered a rise in awareness about this invisible work by showing users their own and other inhabitants' contributions. The multiple ambient displays situated in the participants homes enabled opportunities for reflection and motivated an increased engagement in household tasks.

1 Introduction

The work carried out in the maintenance of households is often considered invisible [1, 2] and has traditionally been a female domain [3]. While improvements towards more gender equality in this domain have been made [4] and a study even found that men are happier when they take on a fair share of household tasks [5], it is still a long way towards closing the gender gap. Beyond family households, shared households of e.g., students co-inhabiting an apartment, also face the challenge of negotiating and sustaining a fair distribution of household chores [6]. In both cases, busy schedules and the lack of objective information about the nature and the extent of the contributions of the different household members impede an assessment of the situation and the coordination of a fair distribution of work.

Making this work visible to the other members of a household through the use of persuasive ubicomp technology could not only provide awareness for the contributions of the individual household members to the common household but also enable a reflection of the current situation and motivate an increased contribution to the household by its less active members. The coordination of household errands has recently been subject to a diary study by Sohn et al. [7], who conclude that design opportunities include mobile tools for the sharing of errands. However this was

S. Berkovsky and J. Freyne (Eds.): PERSUASIVE 2013, LNCS 7822, pp. 180–191, 2013.

reflecting on errands outside of the home; in our case here we are concerned with coordination and contribution to tasks within the home and also exploiting the potential of mobile tools in the form of distributed situated displays to both capture and make visible work contributions. To this end we report on the design and in-situ evaluation of distributed system, Choreflect, to promote the visibility and sharing of household chores.

This paper starts with a presentation of related work and continues with a description of the prototype including the design rationale. After this the study setup and results are outlined. This is followed by a discussion and some concluding remarks. The main contributions of this paper are a field study of a persuasive system exploring the novel approach of multiple distributed and contextually situated ambient displays and a discussion of the emerging design concerns arising from its use. The system further serves as a research vehicle to explore open ended persuasive technologies that foster a reflective approach towards behavior change.

2 Related Work

Ambient and ubiquitous technology denotes the seamless integration of technology into our everyday world and goes back to Weiser's vision of ubiquitous or calm computing [8], further developed in the concept of Ambient Intelligence [9]. The utilization of ambient technology for persuasion enables the presentation of persuasive cues in a contextually adequate way [10], raising the potential for a successful persuasive intervention by utilizing the principle of Kairos, which states that "a computing technology will have greater persuasive power if it offers suggestions at opportune moments." [11].

With the event of smartphones and tablet PCs, a cost efficient platform with integrated networking capacity for the realization of situated and distributed ambient displays [12] has become available, making the deployment of display ecosystems [13] technologically and economically feasible and allowing us to embed such suggestions in the right place at the right time. Our prototype utilizes ambient displays situated in the users' homes to provide context specific feedback and to support task tracking. This approach represents a move on from Mankoff's [14] notion that ambient displays are merely perceived and not used, towards enabling implicit and explicit interaction, in line with Rogers' call for making ubicomp more engaging [15].

HCI research is already exploring the issue of integrating ubicomp applications into the home context. Gaver et al. [16] experimented with the deployment of distributed ubicomp technology in the home in a study of what he terms threshold devices, aimed to connect the home with information from its wider environment. The results showed the potential of such distributed and situated ambient displays and artifacts to engage people in the home. The home context has also been subject to numerous persuasive interventions, mostly in the area of sustainable energy consumption. Systems in this area typically consist of an (ambient) display, usually a single display, that gives feedback about electricity [17] [18], or water consumption [19] with the goal of raising users' awareness about their related behaviors and ultimately influencing them

towards a reduction in energy use. However the input to these systems is generated automatically, while we want to explore the potential of more active input of data (to be discussed later), and the feedback simply serves to show how much energy is consumed by certain actions, such as using the clothes dryer or the dishwasher. Another application area for behavior change systems in the home is health and well-being. Chatterjee et al. [20] used wireless sensors in the home (and on the body) to track activities to support the self management of diabetes via persuasive text messages. While mobile devices are used here, they belong to the person and are not distributed around the home.

The system presented in this paper to promote sharing of household tasks, also relies on the underlying ambient persuasion mechanism [10], i.e. making a hidden factor (energy usage, activities of daily life, contribution to household work) present in this setting visible through the use of mobile and ambient displays. What is novel in this work is using mobile devices as distributed situated persuasive displays rather than having a single device located in a particular location in the home. We also explore the use of active input from the user; even though we could have automated much of the data capture, we wanted to explore if such active involvement can serve as an opportunity for reflection and change.

3 System Prototype

The goal of the Choreflect system is to generate awareness for the contributions made to a common household by its individual members, fostering reflection potentially resulting in a fairer distribution of household work. This section starts with a description of the system from a user perspective and giving the design rationale, and then provides details about the components and technologies that were used for the prototype. The prototype consists of three or four (depending on the requirements of the household) ambient area displays and one master display, each installed in different locations in the household and all connected via internet to the other main component of the system, a web server with a SQL database. Via the distributed system users can see what chores have been done and what needs to be done in a particular room and are able to record tasks they finished. Furthermore, the master display gives an overview about the actions of all inhabitants.

3.1 Using the System

The Choreflect system is distributed in different areas of a home, e.g. the bathroom, the hallway, the kitchen and the living room, which are chosen by the inhabitants themselves (see Fig. 3 later). Each of these areas is equipped with a corresponding area display, which displays the tasks due to be done and the tasks recently completed for that area (see Fig. 1.a and c). With a glance at the device the occupants are able to see whether the area needs attention or not and they also can see which tasks were completed within the last 12 hours.

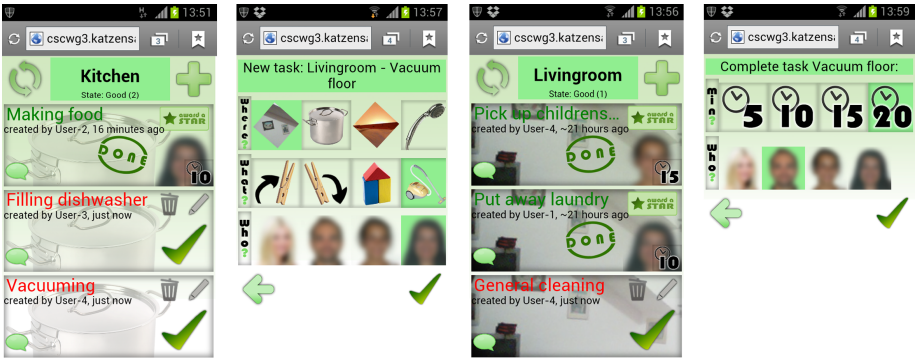


Fig. 1. Area displays: (a) kitchen, (b) new task, (c) living room, (d) complete task

A task consists of several parameters, such as task name, the name of the person who created it and the date of creation. Users are able to create tasks for an area by touching the plus button on the right top of the display. Choreflect lets the users specify the parameters of the new task they intend to create by tapping on the icons for location (default is the area where the device is located), task and user (see Fig.1.b).

Once a task is done, the prototype allows users to mark a task as completed by touching the checkmark of a task representation. A new page appears and the user can pick his/her icon and enter the amount of time it took to complete the task (see Fig.1.d). Completed tasks are visible on the respective area display for 12 hours after completion in order to provide a quick overview of the activity in this area.

The master display (see Fig. 2.), like the area displays, is always active. The master display contains two pages: The overview page is the standard page that is normally displayed on the master display: a colour-coded personalised floor plan on the top indicates the states of the different areas and displays icons for standard tasks. Below, there is a list displaying all the messages of the system. A message represents a user action, e.g. completing or adding a task. The icons on the left side of each message indicate the message type. By tapping on the right side of a done message, users can award a star to another user for the particular task. The display not only shows who received a star for which task, but also by whom it was awarded.

The second page of the master display can be reached by tapping the round bar icon. It gives a complete overview of the efforts the users have invested in the household since they have started using Choreflect. When loading the page, a 'race' starts on the top of the page: The user icons start at the left and move to the right along their lane. The leader of the race (with the icon furthest right) is the person who put in the most effort in cleaning and maintaining the household. The number on the right end indicates the user's percentage of total time invested by all users. Below the race, there are two tables, one for user-based and one for general statistics.

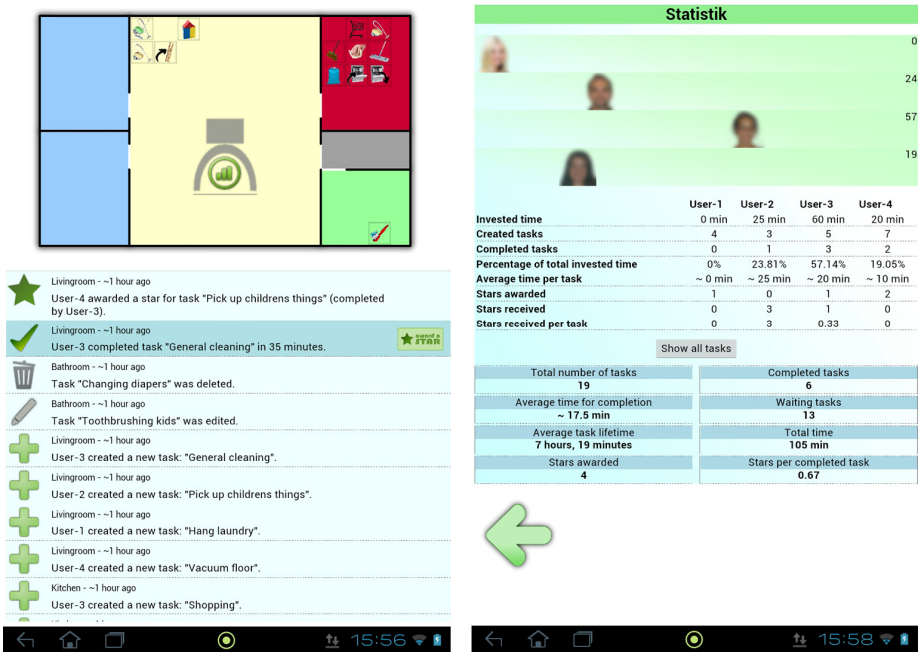


Fig. 2. A master display showing (a) the overview page and (b) the statistics page

3.2 Design Rationale

There are a number of principles underpinning the design of Choreflect: visibility, kairos/situatedness, engagement, recognition, social facilitation, self-monitoring, and rewards. For visibility, the master display is installed at a prominent space that can be seen by everyone multiple times a day. With a glance at the master display users get information about the latest activities. The goal of this permanent display of the actions is to make it much easier for all the inhabitants to see the efforts of others. Likewise, each person's own efforts are more visible, too. The ambient displays are situated in the inhabitants' apartments, providing cues and the possibility for task entry at opportune moments (kairos). Choreflect is designed to engage users in active interaction for entering tasks, with the aim of encouraging them to become more mindful of their contributions. Furthermore, awareness is increased by information displayed at the statistics page, utilizing the persuasive principle of recognition [21].

The prototype's constant display of the efforts invested in the household increases the number of moments where users are confronted with the achievements of others. Relating this information to themselves may lead to an adaption of their behaviour. Oinas-Kukkonen et al. [21] denote this mechanism as social facilitation. According to Fogg [22], this is one of the most powerful persuasion strategies. The display of a race can leverage the related principle of competition [21] between the users, see Fig. 2b.

The prototype lets users easily track their efforts by looking at one of the ambient displays and provides even more specific information at the statistics page of the

master display, enabling a reflection of one's own contribution, a process similar to the principle of self-monitoring by Fogg [11]. Furthermore, the prototype offers different rewards [11], system-triggered and user-triggered. Users can express their appreciation of other users by awarding of stars for tasks. System triggered rewards are:

1. A done message on the master display - The task is also removed from the list of due tasks; 2. A user icon in the task display - After completion, the task disappears, and instead the user's icon is visible - overlaid with the time it took him/her to complete the task; also a 'done stamp' is applied (see Fig. 1a and c); 3. Moving forward in the race - After the completion of the task, the user's icon will advance in the race.

3.3 Technical Description

In the prototype, the devices used as area displays (see Fig. 1) were two Apple iPod Touch (5th generation), three Samsung Nexus Galaxy Smartphones and one Samsung Nexus S. The devices were all WiFi- enabled and always connected to the internet and a power source. A Samsung Galaxy Tab 10.1, a Samsung Galaxy Tab2 10.1 and an Acer Iconia Tab A200, respectively served as the master display (see Fig. 2). An Apache web server with PHP in combination with a mySQL database were used to host the system and create the corresponding html code for the different devices. The software was programmed in PHP, HTML, CSS, Javascript (also jQuery) and mySQL.

4 Study

In order to explore our new approach towards persuasion based on multiple distributed ambient displays, we conducted a 4-week situated field study of the system in two flat shares and two family homes with a total of 17 inhabitants (see Table 1). Only 12 were represented as users in the system however, because the 5 younger children did not use the system explicitly themselves. Prior to the field study, a usability test with three test persons was conducted. The resulting minor issues were addressed. The resulting prototype was further tested by being installed in a researcher's flat (flat share A) so that he could detect and solve early problems. The researcher (A3) was part of using the system like the other inhabitants, but was not included in the semi-structured interviews. As he was part of the first test, he was instead able to conduct a participatory observation throughout the whole test period. The study participants in the flat shares were students and the families were recruited via a local school and neighborhood announcements. The participants were not financially remunerated for their participation.

The overall goal of the study was to explore how such a system would be used in households and whether it could promote better participation in household chores. Thus the study was designed to gain insights regarding interaction with the system, user acceptance, the distributed ambient displays and awareness and reflection. Each four week study with a household started with an initial interview to gather the requirements to customize the system, followed by a deployment phase and final interview. The user interaction with the system was logged and stored in a database.

4.1 Study Setup

At the beginning of the study, the function of the system was explained to the participants and an initial interview about their existing practices regarding household tasks was conducted. This information was then used to adapt the system, which included setting up the users, creating tailored area displays to match the layout of the rooms, defining typical tasks and assigning them to the area displays of the rooms where they are conducted and updating the map in the master display to match the actual floor plan. Together with the users, suitable positions for the main display and the area displays were chosen (see Fig. 3. for some examples). The tablet PC and the smartphones were then fixed in place using removable glue and/or smartphone mounts and tested. After this, the researcher explained the functionality of the system to the users and encouraged them to try out the system for themselves.

Table 1. Summary of households included in the study

Household	No. of displays	Age and Gender	IDs of participating inhabitants
Flat share A	4	25(m), 25(f), 27(m)	A1, A2, A3
Flat share B	4	29(m), 29(m), 22(f), 28(f)	B1, B2, B3, B4
Family C	5	36(f), 37(m), 11(f), 6(f), 1(f)	C1, C2, C3
Family D	5	36(f), 34(m), 6(f), 4(f), 1(m)	D1, D2

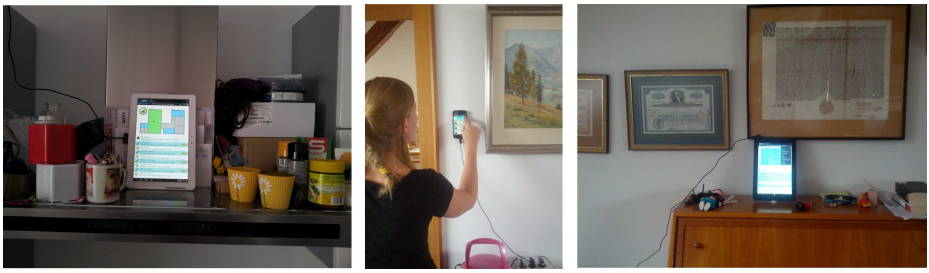


Fig. 3. Choreflect in use: (a) master display in a kitchen, (b) inhabitant interacting with an area display and (c) master display in a living room

At the end of the study the participants were interviewed again. This semi structured interview included questions about their practices regarding household tasks, their interaction with the system and whether they had any suggestions regarding its design. All interviews were recorded and the researcher took detailed notes, using the audio records to fill in missing details after interview. The data of the interview recordings and notes were analyzed qualitatively to get an understanding of use of the prototype and to identify emerging themes, and was complemented with the quantitative data derived from logging the system usage.

5 Results

Based on the log data, during the study Choreflect was used to track a total of 413 tasks and the inhabitants cumulatively recorded spending 8245 minutes in the completion of these tasks. There was a considerable difference in the amount of time invested in the household that was tracked between the flat shares and the families, with family D leading with 6130 min, followed by family C (1525 min), flat share B (345 min) and flat share A (245 min). Overall, this difference hints at the dissimilar amount of work needed to maintain a family household and a flat share and is further explained by the different usage styles for the system: Whereas for the flat shares the system was used to mainly record tasks that were for the collective maintenance of the apartment (e.g. cleaning), the families also included kids related activities. Family D had the widest notion of household activities and also chose to include time consuming tasks like e.g. shopping and children's hobbies. To further illustrate the use of Choreflect and to provide an understanding of its persuasive potential, findings regarding the acceptance of the system, about the distributed ambient displays and the topics of awareness and reflection are presented in the following sections.

5.1 Acceptance of the System

Overall participants liked the system. When asked whether they would like to continue to use the system, the majority of users (9) said, yes, one was neutral and one person said that she didn't want to use the system anymore. Furthermore, in the final interview with the families (C, D) they independently both brought forward the idea of an intermittent use of the system. They suggested that a month seemed like a good time to get an idea about the household tasks carried out in a particular home, so the tracking could be stopped for a while and picked up again when a major change occurs that might influence individual contributions, e.g. school holidays, change of job schedules or a new job. This could help to assess the new situation and support the renegotiation of the individuals' contribution to the common household based on the Choreflect data.

5.2 Distributed Ambient Displays

Over the course of the study, the logs show that each of the displays was actively used by participants for entering and manipulating data. The situated and distributed nature of the displays meant that Choreflect became a part of the participants' home context. As the considerable number of logged tasks, and the results from the interviews show, the displays successfully served as cues in the environment to remind users of the presence of the system and engaged them to use it regularly. For D1, the prototype did well at striking the balance between being there visibly and not being annoying. C1 and C2 observed that the ambient displays throughout their home not only brought the system to mind but also made it easy to create and complete tasks in place. The closest device was use to track the corresponding tasks in an area: " *When I cleaned the kitchen, I mostly used the kitchen device and so on*." (B3).

C2, who was often working late shifts as a baker, said: "*when I came home in the evening and everyone else was sleeping already, I saw the faint glow of the displays and still had a look to quickly check what went on during the day.*" The distributed displays also afforded different interaction styles. Whereas most participants used the master display to get an overview of the current situation in the household, D2 explained that he hardly used the master display (with the exception of awarding stars) and preferred to glance at the area displays to see what was going on. Having multiple displays also added to the robustness of the system, as exemplified by B3's statement: "*One time there was a problem with the kitchen iPod, so I switched to the hallway device*".

5.3 A Tool for Awareness and Reflection

As the study results show, the system enabled reflection on the granularity of single tasks as well as cumulative effort over time, and on an individual as well as on a social level. On an individual level, this can be exemplified by an observation from D1, who pointed out that the system gave her objective information about how long it took to complete a task and that for some tasks this was shorter than she had previously assumed, e.g. emptying the dishwasher only took 5 minutes. Being aware of this also led to a change in behaviour, or as D1 puts it, she decided regarding these tasks that she might as well "*start doing them instead of spending the same five minutes looking at them*". This illustrates that not only the cumulative feedback about the individual contribution over time and the comparison with others, but also the feedback on the granularity level of tasks contains the potential to foster reflection and possibly result in behaviour change.

Reflecting on the amount of work she had done in the household over time, D1, who had described herself as a "lazy housewife" before the prototype was deployed, said that now that she saw objectively how much she was actually doing, she might not be so lazy after all, highlighting a positive change in attitude. Similarly, C1 and C2 both agreed that they knew C1 had a lot of work with the household but did not expect it to be this much, whereas D2 saw confirmed that he was already contributing a considerable amount (45%). In all these cases the system provided opportunities for self-assessment based on objective information.

On a social level, the system also led to an increase in awareness, motivating users towards an increased contribution. As A1 remarked: "*It motivates because you enter your efforts and get points and the others see what you did, officially*". In the playful race competition arising from the comparison between users facilitated by the system, not being last became the goal of several users. In family C, C2 and C3 saw that they did not have a chance for the first place, since C1 was doing far more than they did, but both of them had the desire not to be the one who did the least. This led to both of them doing more for their common household, providing relief for C1, as both she and the pair of them observed. A similar sentiment was voiced by B1, explaining his interpretation of the system: "*I tried to be in a good middle position, that's the average and should be ok, that was my leading thought. Being first doesn't give you an advantage but I didn't want to be last. If everybody is on the same level - that's the*

meaning of the system and this was my goal.". In summary, all but one participants said they felt influenced by the prototype in that it motivated them to do household chores.

The system also influenced the way household tasks where coordinated, according to A1: "*There was no 'you-do-this-and-you-do-that', it was somehow self-regulating.*" and C2, who was very sceptical about his daughter's efforts in the home remarked, in agreement with C1, that after the prototype was introduced she (C3) did her chores without being told so that she can enter them and make them count in the system. The mindful process of entering tasks into the system was an important part of the interaction for the users. When asked whether they would like the system to automatically detect and track the tasks they carried out, all users disagreed, not only because they were sceptical of the technology (privacy aspects and accuracy) but also because they liked entering the tasks, giving them the sense that they completed something now and also served as reward in itself. However, 7 users could still imagine the system automatically generating some tasks that re-occur on a regular basis, such as taking out the garbage, cleaning or vacuuming.

6 Discussion and Conclusion

In this paper we set out to explore the potential of a distributed situated persuasive display in promoting increased sharing of household chores. We designed a system called Choreflect to support a number of specific principles towards resources for reflection and change: working with issues of visibility, situatedness, engagement, recognition, social facilitation, self monitoring and rewards. While designed with these principles in mind, the system was also designed with a view to openness, where its interpretation and use could be appropriated both by individuals within a household and by households themselves, to reflect their specific characteristics.
The data from this study shows that the system was able to be appropriated in different ways and in all cases resulted in increased awareness of the distribution of household chores, fostering the household members to reflect about and reassess their individual contribution. There are a number of lessons to draw out from this for he design of persuasive systems more generally.

The design of Choreflect utilized an ambient persuasion approach [10], which is based on using ubicomp technology to capture a hidden factor in a particular context and make it visible to its inhabitants through ambient displays. Going beyond previous systems in this domain, Choreflect used multiple networked ambient displays situated throughout the household. This proved useful in two ways, on the one hand by enabling a convenient way of tracking tasks through displays in close proximity, and on the other hand by serving as embodied cues in the home environment, constituting a reminder about the presence of the system and giving feedback about the state of the household tasks at opportune moments.

Moreover, the study of the system lends support to the concept of designing open ended persuasive systems ([23]) that enable appropriation through their users based on their individual and collective needs. The prototype allowed this through the initial customization, where the flat shares included only basic household tasks that fell in

the collective responsibility, whereas the families strived to have a more complete mapping of their household activities, which can also be seen from the higher number of completed tasks there. The system also enabled different usage styles, for individual reflection on a task and on cumulative efforts, for social competition, and for the appreciation of one's own and the other's efforts.

Furthermore, the findings regarding the automated vs. manual entry of tasks highlight that in design of PT systems there is a delicate balance between the autonomy of the users and automated acts of the system (cf. [23]). Going beyond this dichotomy, in the case of Choreflect, the manual task entry became an integral part of the interaction process with the system, serving as an embodied statement of one's achievement after completing a task. The system thus helped users to transform the mundane experience of carrying out a household task by enabling them to reward themselves when they successfully finished it, possibly leading to increased motivation [24].

The Choreflect system presented in this paper demonstrated the potential of ambient persuasion systems based on multiple contextually situated displays. It enabled users to become aware of and reflect their own and the other inhabitants' contribution to their common household, thus providing a tool for making this kind of work more visible and potentially leading to its fairer distribution.

References

1. The Invisible Economy and Gender Inequalities: The Importance of Measuring and Valuing Unpaid Work. Pan American Health Organization (2010)
2. Kaplan Daniels, A.: Invisible Work. Social Problems 34 (1987)
3. DeVault, M.L.: Feeding the family: the social organization of caring as gendered work. University of Chicago Press, Chicago (1991)
4. Gershuny, J., Kan, M.Y.: Half-Way to Gender Equality in Work? Evidence from the Multinational Time Use Study. In: Scott, J.L., Dex, S., Plagnol, A. (eds.) Gendered Lives: Gender Inequalities in Production and Reproduction, xiv, 231 p. Edward Elgar, Cheltenham (2012)
5. Scott, J.L., Plagnol, A.: Work–Family Conflict and Well-being in Northern Europe. In: Scott, J.L., Dex, S., Plagnol, A. (eds.) Gendered Lives: Gender Inequalities in Production and Reproduction, xiv, 231 p. Edward Elgar, Cheltenham (2012)
6. Mause, K.: The Tragedy of the Commune: Learning from worst-case scenarios. The Journal of Socio-Economics 37, 308–327 (2008)
7. Sohn, T., Lee, L., Zhang, S., Dearman, D., Truong, K.: An examination of how households share and coordinate the completion of errands. In: Proc. CSCW 2012, pp. 729–738. ACM, Seattle (2012)
8. Weiser, M.: The Computer for the Twenty-First Century. Scientific American, 94–104 (1991)
9. Aarts, E., Marzano, S.: The New Everyday: Views on Ambient Intelligence. 010 Publishers, Rotterdam (2003)
10. 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)

11. Fogg, B.J.: Persuasive Technology: Using Computers to Change What We Think and Do. Morgan Kaufmann Publishers Inc. (2003)
12. Wisneski, C., Ishii, H., Dahley, A., Gorbet, M., Brave, S., Ullmer, B., Yarin, P.: Ambient Displays: Turning Architectural Space into an Interface between People and Digital Information. In: Yuan, F., Konomi, S., Burkhardt, H.-J. (eds.) CoBuild 1998. LNCS, vol. 1370, pp. 22–32. Springer, Heidelberg (1998)
13. Terrenghi, L., Quigley, A., Dix, A.: A taxonomy for and analysis of multi-person-display ecosystems. Personal Ubiquitous Comput. 13, 583–598 (2009)
14. Mankoff, J., Dey, A.K., Hsieh, G., Kientz, J., Lederer, S., Ames, M.: Heuristic evaluation of ambient displays. In: Proc. CHI 2003, pp. 169–176. ACM, Ft. Lauderdale (2003)
15. Rogers, Y.: Moving on from Weiser's Vision of Calm Computing: Engaging UbiComp Experiences. In: Dourish, P., Friday, A. (eds.) UbiComp 2006. LNCS, vol. 4206, pp. 404–421. Springer, Heidelberg (2006)
16. Gaver, W., Boucher, A., Law, A., Pennington, S., Bowers, J., Beaver, J., Humble, J., Kerridge, T., Villar, N., Wilkie, A.: Threshold devices: looking out from the home. In: Proc. CHI 2008, pp. 1429–1438. ACM, Florence (2008)
17. Gustafsson, A., Gyllenswärd, M.: The power-aware cord: energy awareness through ambient information display. In: CHI 2005 EA, pp. 1423–1426. ACM, New York (2005)
18. Kjeldskov, J., Skov, M.B., Paay, J., Pathmanathan, R.: Using mobile phones to support sustainability: a field study of residential electricity consumption. In: Proc. CHI 2012, pp. 2347–2356. ACM, Austin (2012)
19. Erickson, T., Podlaseck, M., Sahu, S., Dai, J.D., Chao, T., Naphade, M.: The dubuque water portal: evaluation of the uptake, use and impact of residential water consumption feedback. In: Proc. CHI 2012, pp. 675–684. ACM, Austin (2012)
20. Chatterjee, S., Byun, J., Pottathil, A., Moore, M.N., Dutta, K., Xie, H(Q.): Persuasive Sensing: A Novel In-Home Monitoring Technology to Assist Elderly Adult Diabetic Patients. In: Bang, M., Ragnemalm, E.L. (eds.) PERSUASIVE 2012. LNCS, vol. 7284, pp. 31–42. Springer, Heidelberg (2012)
21. Oinas-Kukkonen, H., Harjumaa, M.: Persuasive Systems Design: Key Issues, Process Model, and System Features. Communications of the Association for Information Systems 24 (2009)
22. Fogg, B.J.: The Six Most Powerful Persuasion Strategies. In: IJsselsteijn, W.A., de Kort, Y.A.W., Midden, C., Eggen, B., van den Hoven, E. (eds.) PERSUASIVE 2006. LNCS, vol. 3962, p. 6. Springer, Heidelberg (2006)
23. Reitberger, W., Güldenpfennig, F., Fitzpatrick, G.: Persuasive Technology Considered Harmful? An Exploration of Design Concerns through the TV Companion. In: Bang, M., Ragnemalm, E.L. (eds.) PERSUASIVE 2012. LNCS, vol. 7284, pp. 239–250. Springer, Heidelberg (2012)
24. Esch, T., Stefano, G.B.: The neurobiology of pleasure, reward processes, addiction and their health implications. Neuroendocrinology Letters 25, 235–251 (2004)

I Didn't Know That Virtual Agent Was Angry at Me: Investigating Effects of Gaze Direction on Emotion Recognition and Evaluation

Peter A.M. Ruijten, Cees J.H. Midden, and Jaap Ham

Eindhoven University of Technology, The Netherlands

Abstract. Previous research has shown a link between gazing behavior and type of emotion felt. It appears that approach-oriented emotions are better perceived in combination with a direct gaze, whereas avoidance-oriented emotions are better perceived in combination with an averted gaze. In this study, we investigate whether this effect can be applied to persuasive social agents. We hypothesized that an approach-oriented emotion is more credible when combined with a direct gaze, whereas an avoidance-oriented emotion is more credible when combined with an averted gaze. This was tested with both an implicit categorization task and an explicit evaluation. The hypothesis was supported for angry expressions, but not for sad ones. Implications for further research and the design of effective persuasive agents are discussed.

1 Introduction

Imagine having a virtual agent that assists you in your daily life. It provides tips on how you could save energy at home, it manages your agenda and it manages your personal security system. This agent uses persuasion techniques that are known from human-human interactions. Ultimately, the agent becomes like a companion to you, partly because it uses emotional expressions in its communication. Would this also happen if the agent's emotional expressions were ambiguous? In this paper, a crucial cue of a successful human-agent interaction, gazing behavior, is investigated.

In general, people show a tendency to respond socially to technology [1]. When an interactive social agent or even a boxed computer greets us, we respond to this message in a similar way as when another person greets us. These effects could be explained by the notion that interactive social agents make people feel like interacting with a social source [2].

So, increasing the perceived socialness of an interactive social agent or a computer interface could increase the effects it has on people's attitudes and related behavior. To be effective in influencing behavior, a persuasive agent should therefore at least be believable. As Bates denotes, appropriate timing and clear expression of *emotions* is a central requirement for agents to be believable [3]. He also states that: "Emotion is one of the primary means to achieve this believability, this illusion of life, because it helps us know that characters really care about what happens in the world, that they truly have desires." [3, p. 6]. To achieve this illusion of life, or credibility, all that is

S. Berkovsky and J. Freyne (Eds.): PERSUASIVE 2013, LNCS 7822, pp. 192–197, 2013.

needed is a successful social interaction. Even when people are explicitly made aware of the fact that information given by a social agent is randomly chosen, they still change their behavior according to this information [4]. Before an effective interactive social agent can be successfully designed, an understanding of social interactions between people is crucial.

One issue that plays a critical role here is our capacity of processing gaze information [5]. The use of gazing is often related to the emotion that is felt. People gaze more while they are seeking friendship or when they communicate threat and dominance during conversations [6]. In contrast, people gaze less as a result of heightened anxiety or increased depression [6]. This indicates that there is a connection between type of emotion felt and gazing behavior.

In a series of studies, Adams and Kleck investigated this interaction between gaze direction and facial expressions of emotion in human perception. They stated that, "As a social signal, direction of gaze has been implicated in signaling an expressor's approach-avoidance behavioral tendencies" [7, p. 3]. It was investigated if direct gaze increases the perception of *approach*-oriented emotions like anger and joy, whereas averted gaze increases the perception of *avoidance*-oriented emotions like fear and sadness. Results suggested that, in the perception of human emotions, anger and joy expressions were indeed more quickly recognized when presented with direct gaze, whereas fear and sadness expressions were more quickly recognized when presented with averted gaze [8].

In the current study, this interaction between gaze direction and facial expressions will be investigated in the domain of social agents. We argue that people unconsciously ascribe socialness to these agents, and that this process causes responses similar to those found by Adams and Kleck. That is, we hypothesize that an *approach*-oriented emotional expression (anger) of a social agent will be more quickly recognized and evaluated as being more credible when combined with a *direct* gaze than with an averted gaze, and that an *avoidance*-oriented emotional expression (sadness) of a social agent will be more quickly recognized and evaluated as being more credible when combined with an *averted* gaze than with a direct gaze.

2 Methods

2.1 Participants and Design

Twenty-five male and 15 female students (ages between 18 and 26) were recruited. Participants received either course credit or €3 for their participation.

The study consisted of three parts. In the first part, participants performed an emotion recognition task. This task was adapted from [8] and had a 2 (emotional expression: anger vs. sad) X 2 (gaze direction: direct vs. averted) within-subjects design. In the second part, we aimed to explore if a relation exists between emotion recognition of virtual agents and implicit evaluations of robotic agents. For this we used an implicit association test [9], in which associations of robots with socialness were measured. In the third part, participants answered questions about the credibility of the expressions of the virtual agents used in the emotion recognition task.

2.2 Materials

Emotion Recognition Task. In the emotion recognition task, participants were shown pictures of virtual agents, and had to indicate whether the expressed emotion was either anger or sadness. Pictures of four female and four male virtual agents were used. Eight different pictures were generated for each agent. Half of them contained a sad expression and the other half an angry one. Because each expression was displayed twice in averted gaze (left and right), each was also presented twice in direct gaze to balance out the design. All pictures were used twice, and the total of 128 pictures were presented in random order. All expressions were categorized either angry or sad by pressing the 'A'- or 'L'-key. Labels were counterbalanced to control for dominance effects. The dependent variables were response times and number of errors made.

Implicit Association Test. The Implicit Association Test (IAT) followed a standard IAT procedure [9]. The target concepts were robots and men, the target attributes were social and nonsocial, and the variable that was measured is response latency. Past research has shown that this variable can be interpreted as a measure of implicit attitude towards the target-concepts [9]. We refer to trials where 'social' and 'men' are categorized under one key as congruent, and trials where 'social' and 'robot' are categorized under one key as incongruent.

Explicit Evaluations. To measure credibility, participants evaluated the virtual agents' expressions on intensity, realism, and human-likeness scales ranging from 1 (not at all) to 7 (extreme). Each agent was used once in this questionnaire, leading to a total of 24 questions. The congruent C and incongruent (sad-direct and anger-averted) expression-gaze combinations were equally distributed. For the dependent variables, all twelve questions on both the congruent (α = 0.76) and incongruent (α = 0.79) expressions were averaged.

2.3 Procedure

Upon entering the lab, participants first performed the emotion recognition task. In this task, the expressions of virtual agents were categorized as being angry or sad. We expected angry expressions to be recognized more quickly when combined with a direct gaze, and, in contrast, sad expressions to be recognized more quickly when combined with an averted gaze.

Next, participants completed the IAT. This task was initially included as a filler task irrelevant to our current hypothesis. It is intended to measure implicit associations between robots and socialness.

Finally, participants evaluated the pictures of emotional expressions that were used in the emotion recognition task. After finishing this questionnaire, participants indicated age and gender and left the room. They were debriefed, paid, and thanked for their contribution.

3 Results

Prior to analyses, incorrect responses in the emotion recognition task (8.7%) were replaced by the mean response time of all responses. This method was adapted from [10]. Another variable was created to indicate the number of errors a participant made. All response times' data were log-transformed. For ease of interpretation, we converted data back into milliseconds for reporting means. One participant had response times slower than three standard deviations from the mean and was excluded from further analyses.

3.1 Emotion Recognition Task

To test the main hypothesis, the mean response times were submitted to a 2 (emotion: angry vs. sad) X 2 (gaze direction: direct vs. averted) ANOVA. A significant main effect of emotion emerged, $F(1, 38) = 24.17$, $p < 0.001$, $\eta_p^2 = 0.39$. Sad expressions ($M = 1124$, $SE = 42.63$) were recognized more quickly than angry expressions ($M = 1243$, $SE = 45.68$). Furthermore, a main effect of gaze emerged, $F(1, 38) = 11.04$, $p = 0.002$, $\eta_p^2 = 0.23$. Expressions with a direct gaze ($M = 1152$, $SE = 40.97$) were recognized more quickly than expressions with an averted gaze ($M = 1215$, $SE = 45.32$). These main effects were qualified by the predicted interaction between emotion and gaze direction, $F(1, 38) = 7.18$, $p = 0.011$, $\eta_p^2 = 0.16$. More specifically, the congruent (sad-averted and anger-direct) expressions ($M = 1158$, $SE = 41.00$) were recognized more quickly than the incongruent (sad-direct and anger-averted) expressions ($M = 1209$, $SE = 45.18$).

Similar results were found on the number of errors made in the task. To test the hypothesis, the number of errors made were submitted to a 2 (emotion: angry vs. sad) X 2 (gaze direction: direct vs. averted) ANOVA. A significant main effect of emotion emerged, $F(1, 38) = 4.37$, $p = 0.043$, $\eta_p^2 = 0.10$. Angry expressions ($M = 6.46$, $SE = 0.84$) were falsely recognized more often than sad expressions ($M = 4.67$, $SE = 0.65$). Furthermore, a main effect of gaze emerged, $F(1, 38) = 9.10$, $p = 0.00$, $\eta_p^2 = 0.19$. Expressions with an averted gaze ($M = 6.33$, $SE = 0.68$) were falsely recognized more often than expressions with a direct gaze ($M = 4.79$, $SE = 0.65$). These main effects were qualified by the predicted interaction between emotion and gaze direction, $F(1, 38) = 7.17$, $p = 0.011$, $\eta_p^2 = 0.16$. More specifically, the incongruent (sad-direct and anger-averted) expressions ($M = 4.74$, $SE = 0.56$) were falsely recognized more often than the congruent (sad-averted and anger-direct) expressions ($M = 6.38$, $SE = 0.79$).

3.2 Explicit Evaluations

To test the credibility of congruent expressions versus incongruent expressions, the averaged responses on all questions for these two groups were submitted to a paired samples t-test. As expected, results indicated a significant difference, $t(38) = 6.93$, $p < 0.001$, $r = 0.75$. Congruent expressions ($M = 4.52$, $SE = 0.11$) were rated more

credible than incongruent expressions (M = 3.88, SE = 0.12). Furthermore, the difference between credibility of congruent and incongruent expressions was positively correlated with the difference in response times of congruent and incongruent expressions in the emotion recognition task, $p = 0.048$, $r = .27$.

3.3 Implicit Association

To test if a main IAT-effect occurred, mean response times on congruent (man-social) and incongruent (robot-social) blocks were submitted to a paired samples t-test. As expected, a main effect was found, $t(38) = -5.90$, $p < 0.001$, $r = 0.69$. Responses on congruent trials (M = 689, SE = 16.78) were faster than responses on incongruent trials (M = 835, SE = 31.30). Thus, participants were faster in combining congruent categories (humans with social words versus robots with nonsocial words) than in combining incongruent categories (humans with nonsocial words versus robots with social words). These results will be used to further develop this measure.

Interestingly, the difference in response times between congruent (man-social) and incongruent (robot-social) trials was positively correlated with the difference in errors made between the congruent (sad-averted and anger-direct) and incongruent (sad-direct and anger-averted) expressions in the emotion recognition task, $r = 0.33$, $p = 0.038$.

4 Discussion

In this study, we investigated the interaction between gaze direction and emotion recognition in the domain of social agents. It was argued that people unconsciously ascribe socialness to these agents, and that this process causes differences in the recognition and evaluation of emotional expressions with direct and averted gaze. In general, the data confirmed our expectations. When the gaze direction matched the motivational orientation belonging to a specific type of emotion, this emotion is recognized faster and evaluated more credible. These results conceptually replicate earlier findings on emotion recognition in humans [7][8]. The same pattern of results was found in explicit evaluation of the social agents. This suggests that matching gaze direction to type of emotion not only positively influences recognition of this emotion, but it also increases the agent's credibility. This enhanced credibility could potentially increase the persuasiveness of social agents.

Responses of our participants in the emotion recognition task were on average much slower than responses found by Adams and Kleck [8]. This could be caused by several reasons. First, they used pictures of real humans and not artificial agents. It can be argued that emotion recognition in humans is faster than emotion recognition in virtual representations of humans, because they do not yet perfectly represent human emotions. Also, in [8] they used a fixation point that marked the position of the presented stimulus to increase the focus of participants on the screen. We did not include such a fixation point. Despite of these differences, we still find patterns of results similar to earlier findings.

Summarizing, this study extends earlier work on emotion recognition into the domain of social agents. Our findings largely confirmed the expectations. In the design of social agents that are intended to interact with human users, it is important to match gaze behavior to expressed emotions. Further research is needed to include other approach- and avoidance-oriented emotions (like joy and fear). Furthermore, research is needed to investigate if an agent indeed becomes more *persuasive* when its gaze behavior matches its conveyed emotions. Doing this will increase the agent's credibility, and, ultimately, increase the persuasive strength of such agents.

References

1. Reeves, B., Nass, C.I.: The media equation: How people treat computers, television and new media like real people and places. Cambridge University Press, New York (1996)
2. Ruijten, P.A.M., de Kort, Y.A.W., Kosnar, P.: Bridging the Gap between the Home and the Lab: A Qualitative Study of Acceptance of an Avatar Feedback System. In: Bang, M., Ragnemalm, E.L. (eds.) PERSUASIVE 2012. LNCS, vol. 7284, pp. 251–255. Springer, Heidelberg (2012)
3. Bates, J.: The role of emotion in believable agents. Communications of the ACM 37(7), 122–125 (1994)
4. Midden, C., Ham, J.: The Illusion of Agency: The Influence of the Agency of an Artificial Agent on Its Persuasive Power. In: Bang, M., Ragnemalm, E.L. (eds.) PERSUASIVE 2012. LNCS, vol. 7284, pp. 90–99. Springer, Heidelberg (2012)
5. Baron-Cohen, S.: Mindblindness: An essay on autism and theory of mind. MIT Press (1997)
6. Kleinke, C.L.: Gaze and eye contact: A research review. Psychological Bulletin 100(1), 78–100 (1986)
7. Adams, R.B., Kleck, R.E.: Effects of direct and averted gaze on the perception of facially communicated emotion. Emotion 5(1), 3–11 (2005)
8. Adams, R.B., Kleck, R.E.: Perceived gaze direction and the processing of facial displays of emotion. Psychological Science 14(6), 644–647 (2003)
9. Greenwald, A.G., McGhee, D.E., Schwartz, J.L.: Measuring individual differences in implicit cognition: The implicit association test. Journal of Personality and Social Psychology 74(6), 1464–1480 (1998)
10. Greenwald, A.G., Nosek, B.A., Banaji, M.R.: Understanding and using the implicit association test: I. An improved scoring algorithm. Journal of Personality and Social Psychology 85(2), 197–216 (2003)

Three Themes for Designing Games That Aim to Promote a Positive Body Perception in Hospitalized Children

Ruth Sancho Huerga, Jennifer Lade, and Florian 'Floyd' Mueller

Exertion Games Lab, RMIT University, Melbourne, Australia
{ruth,jennifer,floyd}@exertiongameslab.org

Abstract. Hospitalized children often experience physical changes that negatively affect their bodily perceptions, thereby adding to the stress of being sick. Existing approaches to supporting hospitalized children such as those promoted by the Clown Doctors use play to distract the child from negative bodily perceptions. In contrast, we propose reframing the bodily perception of these children through bodily virtual play facilitated by their imagination. We explore this design space through an analysis of the literature combined with design explorations around play and bodily imagination. This research results in a set of themes for games that aim to reframe bodily perception to a more positive self-image full of creative potential. We envisage that our work could help designers who aim to create digital play for sick children.

1 Introduction

Children's hospitals are increasingly introducing complementary care programs that facilitate play for sick children. However, we note that there has been little work in designing play projects that reframe body perception in hospitalized children. In this context our work utilizes digital and physical play to improve children's health. We propose a design space with a set of three themes that can be used as a guide for games designers when designing play that aims to reframe sick children's body perception.

Whilst conducting a review of the literature and engaging with the design process, we worked as a team that included academics, theatre and bodily expression experts, designers and computer programmers. We also engaged two children so far during the game design explorations.

In this paper we present an overview of current works and literature that are relevant for creating games for children in hospital. We present concepts relevant to creating games for health and discuss our approach when creating games for sick children focused on reframing bodily perception.

2 Related Work

Over the last 20 years different therapeutic play works addressing body perception in sick children have been developed to be used as mental and physical complementary care in hospitals. Associated studies have found that this alternative non-medical

S. Berkovsky and J. Freyne (Eds.): PERSUASIVE 2013, LNCS 7822, pp. 198–203, 2013.

treatment can help sick children through therapeutic play to enhance their coping, development and adaptation to life changes, supporting complementary care for mental and physical health [3]. For instance, organizations such as Clown Doctors [8] and Therapeutic Clowns [9] address the issue of body changes and body perception through clown improvisation. Clowns use props such as puppets, music and magic to help them approach and express feelings about their changing body perception in a safe, comfortable and creative environment. Studies assert that clowns performing with children requiring surgery can be effective for managing children's anxiety [10]. However, there are no studies articulating if and how such approaches can facilitate a positive body perception. We therefore see an opportunity to use a non-medical care method such as therapeutic play to support the body perception of sick children.

Children find digital technology a very familiar media [6] and engagement in this media allows them to explore their imagination [2]. We therefore also focus on the digital to explore imagination. Research previously suggested that combining digital media with physical objects could be used for physical recovery treatment [11]. For example "Elements" by Jonathan Duckworth et al. [11] is an interactive table top that supports rehabilitation of patients recovering from traumatic brain injury. Elements supports painting, creating music, matching objects and shapes by using physical objects in a digital environment to enrich the experience of recovery treatment. This inspired us to believe that physical objects could be useful for physical recovery treatment while at the same time a digital environment could create a playful platform to apply therapeutic exercises. The "Elements" system targets body perception in both areas, mental and physical health and facilitates imagination, however it is not addressed at sick children.

Digital objects have recently been designed to support complementary care within a therapy context. Magic Land [6] is an interactive table top that consists of four different toys implemented on a smart table to help children understand and overcome feelings of anxiety. The following games demonstrate how the system enables a wide range of different activities to support complementary care: "Flying Feathers" supports imagination, "Rosebush" supports play, "Hero/Avatar" supports fantasy playing and "Water" supports new play opportunities. This suggests to us that digital interactive play can support therapeutic play. However, Magic Land does not address to physical health so far.

In summary, we have presented a number of projects that target sick children while considering body perception. However, we find a gap when it comes to complementary care that aims to help children developing different body perception. We see an opportunity to explore this gap through the development of a design space and we focus on facilitating this through a set of themes for games that aim to reframe bodily perception.

3 Design Approach

We used an examination of existing literature and design explorations to derive a design space for designers to create games that could reframe sick children's body perception. We identified three main themes for reframing body perception. These three themes were derived from analyzing papers selected from these domains: health,

complementary care, games, children's psychology and body perception. We further refined these three themes by exploring designs around physical play, virtual play and bodily – virtual play.

Bodily Play. We created props such as puppets made out of medical gloves and puppets made out of x-rays of a child. During this development and associated play we noticed that the child was mirroring his puppet's movements and behaviors. He imagined he was a lion and started physically embodying the lions' power through his body, moving his arms and hands up and down as claws, roaring, attacking and feeling empowered [fig. 1]. The child reframed his bodily perception through bodily play transforming his body into a new one full of power.

Fig. 1. Child mirroring the puppet's movements and behaviors

Virtual Play. We created "Laughter Tree", a Kinect game based on a laughter therapy activity [1]. We developed this game to explore bodily virtual play to reframe body perception. We designed a tree made of pictures of the children's arms and hands that drops leaves which laugh when the player grabs them via the avatar [fig.2]. The game has so far been played by seven adults and two children, who reported the game to be fun and engaging.

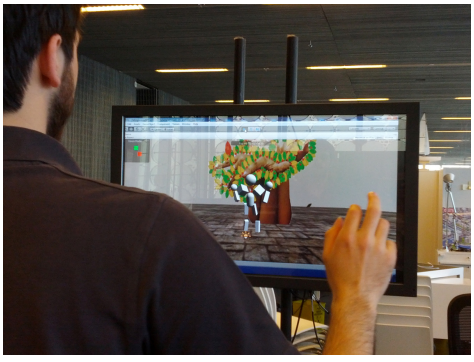


Fig. 2. "Laughter Tree": moving one's arm facilitates laughter triggered in the virtual world

Bodily – Virtual Play. Through our design explorations we were inspired to combine the two approaches, for example, we were interested in using the glove puppets in a digital environment, effectively using them as controllers [fig.3]. Using glove puppets as controllers could result in a combined physical and virtual system that facilitates engagement further [2].

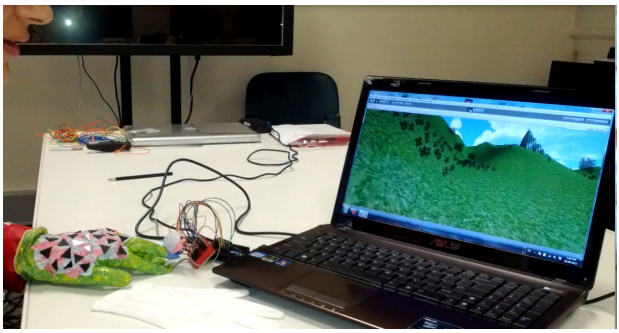


Fig. 3. Puppets made out of medical gloves to control virtual environments

4 Design Space for Reframing Body Perception through Bodily Virtual Play

After reviewing the literature and analyzing the design explorations we came up with three main themes that we believe are a good start for designing the reframing of body perception through bodily virtual play.

4.1 Empowerment of the Body through Bodily Play

Children in hospital environments can feel restricted, stressed and vulnerable [4]. We believe that setting a scenario where the child controls the situation and feels ownership of the environment can change this. We propose that through bodily play we can empower the child's perception of his/her body. Through bodily play the sick child may reframe her body perception while using it for an entertaining, positive and constructive purpose rather than addressing it only for medical reasons.

4.2 Engagement of the Body through the Imagination Facilitated by Digital Media

Children play because they have fun, enjoy, laugh and develop their imagination [5]. It is through the imagination that children can learn in the most constructive, positive and creative way [5]. The imagination is enriched by the use of digital media [3]. Through the development of Laughter Tree while using Jacques L'ecoq's 'outercourse' improvisation techniques [7] we found out that the child can imagine better ways of using her body when facilitated by the use of digital media. We believe

facilitating engagement of the body through the imagination facilitated by digital play enhances the change from the perception of a sick body to a different one full of creative potential.

4.3 Sociability through Reframed Body Perception Facilitated by Digital Media

Hospitalized children can experience adverse side effects in their body due to medical treatment. These changes affect children's mental and physical body perception resulting in a decrease of children's desire to socialize [4]. Those children isolate themselves avoiding social contact that ends up in a deterioration of their mental and physical health. We believe that designing games that reframe the body perception can help sick children to empathize with other sick children and feel socially supported rather than isolated and lonely. During the design process we discovered that enabling multi-players functionality could not only contribute to engagement, but could also help sick children to socialize with other sick children.

5 Discussion and Conclusion

The design space and the three themes have yet to be confirmed. Although we have presented the games to adult participants and two children, we have not yet presented it to sick children. Our next step is to undertake a study with sick children in hospital to confirm the design space and the three themes.

In this paper we have presented our design explorations combined with an analysis of the literature resulting in the creation of a design space and a set of three themes for creating games that aim to reframe bodily perception. We anticipate that our work can help designers to use digital play for positive health change in sick children.

Acknowledgments. We would like to thank Rhys Sullivan, Mathew Sullivan, Christopher Mackenzie, Yueyue Mo, Courtney Blackney, Leon Eduardo Diaz Jacobi and Arrulfo Diaz Trujillo for their collaboration.

References

1. Adams, R.E., McGuire, A.F.: Is Laughter the Best Medicine? A Study of the Effects of Humor on Perceived Pain and Affect. Activities, Adaptation and Aging 8, 157–175 (1986)
2. Calvillo-Gamez, E.H., Cairns, P.: Pulling the Strings, A Theory of Puppetry for the Gaming Experience. In: Proc. of the Philosophy of Computer Games, pp. 308–323. University Press, Potsdam (2008)
3. Koller, D., Gryski, C.: The Life Threatened Child and the Life Enhancing Clown: Towards a Model of Therapeutic Clowning. Evidence-Based Complementary and Alternative Medicine 5, 17–25 (2008)
4. Larouche, S.S., Chin-Peuckert, L.: Changes in Body Image Experienced by Adolescents with Cancer. Journal of Pediatric Oncology Nursing 23, 200–209 (2006)

5. Ott, M., Pozzi, F.: Digital games as creativity enablers for children. Behaviour & Information Technology 31, 1011–1019 (2012)
6. Pykhtina, O., Balaam, M., Wood, G., Pattison, S., Kharrufa, A., Olivier, P.: Magic Land: The Design and Evaluation of an Interactive Tabletop Supporting Therapeutic Play with Children. In: Proc. of the Designing Interactive Systems Conference, pp. 136–145 (2012)
7. Sherman, J.F.: The Practice of Astonishment: Devising, Phenomenology, and Jacques Lecoq. Theatre Topics Journal 20, 89–99 (2010)
8. The Humour Foundation, Clown Doctors, http://www.humourfoundation.com.au/our-work/clown-doctors.html
9. The Canadian Association of Therapeutic Clowns, http://www.therapeuticclowns.ca/
10. Vagnoli, L., Caprilli, S., Robiglio, A., Messeri, A.: Clown Doctors as a Treatment for Preoperative Anxiety in Children: A Randomized, Prospective Study. Pediatrics 116, 563–567 (2005)
11. Wilson, P., Duckworth, J., Munford, N., Eldridge, R., Guglielmetti, M., Thomas, P., Shum, D., Rudolph, H.: A Virtual tabletop workspace for the assessment of upper limb function in Traumatic Brain Injury (TBI). In: Proc. of Virtual Rehabilitation, pp. 14–19. IEEE (2007)

Persuading Consumers to Reduce Their Consumption of Electricity in the Home

Alan F. Smeaton[1] and Aiden R. Doherty[1,2]

[1] CLARITY: Centre for Sensor Web Technologies and School of Computing, Dublin City University, Glasnevin, Dublin 9, Ireland
[2] British Heart Foundation Health Promotion Research Group University of Oxford, U.K.
alan.smeaton@dcu.ie

Abstract. Previous work has identified that providing real time feedback or interventions to consumers can persuade consumers to change behaviour and reduce domestic electricity consumption. However, little work has investigated what exactly those feedback mechanisms should be. Most past work is based on an in-home display unit, possibly complemented by lower tariffs and delayed use of non-essential home appliances such as washing machines. In this paper we focus on four methods for real time feedback on domestic energy use, developed to gauge the impact on energy consumption in homes. Their feasibility had been tested using an experimental setup of 24 households collecting minute-by-minute electricity consumption data readings over a period of 18 months. Initial results are mixed, and point to the difficulties of sustaining a reduction in energy consumption, i.e. persuading consumers to change their behaviour. Some of the methods we used exploit small group social dynamics whereby people want to conform to social norms within groups they identify with. It may be that a variety of feedback mechanisms and interventions are needed in order to sustain user interest.

1 Introduction

The topic of real-time electricity monitoring in domestic environments has recently attracted attention from a range of research communities. Much of the interest in this area is based around smart meters which allow generation of data for per-minute or even finer-grained sampling of domestic energy use. Such monitoring can generate huge streams of data and once usage data is sampled it can be uploaded to utility providers for monitoring and optimisation of the overall power network. Usage data can also be made available to customers directly so they can see their own usage in real time. This offers the potential for incentivising customers to reduce their consumption during the peak demand times, typically between 5pm and 7pm on weekdays. The motivation for this is that utility providers have to build and operate enough power generation capabilities to cover peak demands, and if smart metering can incentivise energy

S. Berkovsky and J. Freyne (Eds.): PERSUASIVE 2013, LNCS 7822, pp. 204–215, 2013.

customers to reduce their load during those peak times then less power generation capacity is required, overall costs are reduced, carbon footprints are smaller, and everybody is a winner.

For most of the work in this area, real time usage notification is via an in-home display (IHD) unit, a small device which is typically graphical in interface, with a numeric display and is usually a mono-chromatic LCD. Yet simply deploying smart meters and IHDs falls short of inducing behaviour change. In this paper we report a series of tests on our own deployment. The number of subjects in our study are small compared to others but our work is different from most studies because we exploit social grouping as part of the incentivisation. The rest of this paper is organised as follows, In the next section we briefly review related work also reporting user trials and studies. We then describe a recently-published study reporting evidence from two large deployments, followed by a description of our four feedback mechanisms and then an analysis as to their efficacy in promoting change in behaviour. We conclude the paper with recommendations and pointers to future work.

2 Related Work: Studies on Energy Usage

One of the earliest and most ambitious examples of a smart metering trial is the Olympic Peninsula study on 112 residential homes which were fitted with smart appliances (which could be switched on/off remotely), including heating thermostats, water heaters and clothes dryers. This followed previous work [11] where usage of a single appliance, a domestic washing machine, was the basis for various energy usage feedback to users. In the Olympic Peninsula study, data was uploaded to a central server every 5 minutes with price-plans changing in real-time based on the energy demand requested across the 112 residences, so the cost of energy was spread across the residences. On average, consumers saved approximately 10% on their bills from the previous year and from a power generation point of view the peak distribution load requests were significantly reduced [18]. However one drawback with this project was the need to install sensors and actuators on each plug in each home which is an expensive undertaking.

Another study showing the benefits of real-time feedback by Dada *et al.* [3] demonstrated that providing energy consumption information to users can help in deciding when to switch on/off household appliances, pointing out the benefits of real-time information as opposed to static, post-hoc carbon calculators. Recently the Irish Commission for Energy Regulation published results from an 18-month smart metering trial on a representative sample of 5,028 residential consumers which included Time-of-Use (ToU) tariffs and customer-side management stimuli such as a detailed bill with explanations and a real time usage monitor. The findings were that 82% of participants made some change to the way they use electricity as a result of knowing more about their energy consumption and the average overall reduction in use was 2.5% with demand at peak times reducing by 8.8% [10]. Other smart meter' trials have similarly shown that

increased awareness of energy use, and helping consumers interpret that data, consistently leads to reductions and flattening of the peak demand curve which can lead to a typical saving of 5-15% [4,19] or 5-12% [5] and a summary of the work done in this area can be found in Hazas *et al.* [8].

The results of these and other findings have convinced many major electricity providers to roll out smart meter programs, e.g. the main Italian energy provider, ENEL, has already installed about 30 million metering units [14] and according to an EU single market electricity directive (2009/72/EC) at least 80% of Europe's consumers must be equipped with smart meters by 2020, subject to positive economic assessment of all the long-term cost and benefits.

The above evidence suggests that smart metering can play a vital role in helping electricity generation companies reduce their carbon footprints through better predicting overall electricity demand and also through customers consuming less energy or less energy at times of peak demand. However what has not been established are the types of methods to actually inform customers of their energy use [6]. Fischer comments that *"... implementation of feedback is lagging way behind knowledge ..."* [5].

From the design point of view there is a general agreement as to the criteria required specifically when informing users of their domestic energy usage, namely:

- give real-time feedback [2,16]
- include colour coding [20,13]
- build in a comparison to baselines & contextual information [2,13,16]
- relate usage to actual costs [2,13]
- provide self-comparison [16,20]
- include good asthetics [9,16]
- ensure availability on mobile devices, anywhere, any time [13,20]
- include easy access to historical data [16,20]

However, while some of the work referenced in the guidelines above has produced prototype systems incorporating many of the aforementioned considerations, they have not yet been evaluated by participants in user studies [13,20] and most large-scale user trials or studies have used the simple IHD at best, as the mechanism to feedback usage information. Clearly we should focus on mechanisms to change participant habits/behaviour rather than merely report electricity usage [15,17].

In this paper we discuss our own design and experimental setup experiences where we have gathered *minute-by-minute* electricity usage data from 24 households over an 18-month period and we introduced 4 different interventions to try to induce a change in behaviour. The lessons we learned are important as we will see later on. but before introducing our own work, in the next section we will summarise other recently-published work reporting two large field trials.

3 Social Norms for Incentivising Behaviour Change in Energy Consumption

One of the most significant experiments in recent times in terms of exploring what influences human behaviour, was the work by Goldstein *et al.* [7] on increasing the rate at which guests recycle bathroom towels in hotels. This experiment used a variety of messages in hotel rooms to encourage guests to recycle their towels rather than throwing them on the floor thereby indicating a need for clean ones when their room was made up next day. Their results found that by using messages which introduced the guests into social groups they thereby exploited our natural tendency to follow the behaviour of others with similar characteristics to ourselves. By using simple messages with text like *"other guests who stayed in this room ..."*, or *"join the men and women who are helping ..."* rather than the usual *"help save the environment"* message, which they refer to as provincial norms, they observed a significant increase in participation in the recycling program.

This idea of binding participants to form some social grouping was exploited on a much larger scale in a series of recent trials on 170,000 domestic energy consumers, reported in [1]. Two studies in the Sacramento Municipal Utility District of California and the Puget Sound area of Washington state, provided feedback to customers on their home electricity and natural gas usage, with the novelty being a focus on peer comparisons. Customers were provided with monthly or quarterly feedback reports which compared their usage to those of their neighbours in similarly-sized houses and these formed the peer groups.

The net outcome of the peer comparison intervention was that households in receipt of the peer comparison reports made significant and lasting reduction in their energy usage, so much so that if the savings were extrapolated to the whole of the city of Sacramento then energy bills would be reduced by over $15M per year [1]. There are of course caveats with these results, such as the fact that the focus has been on single family, detached homes at the higher end of the market rather than on apartments or in areas of less affluence, but still the findings are significant in that they show peer comparison feedback can yield sustained behavioural change and can be scaled.

The theory of small-group peer comparison tells us that people tend to form bonds with smaller groups with which they feel more connected and which turns into higher trust values so they are more likely to conform to the norm in that group. Other work on studies of peer comparison has tended to reveal the "boomerang" effect where those with less-than-normal usage then tend to increase their consumption, as well as the free riders who tend to emerge when we find out that everybody else is saving. This theory of conformance to social norms has been espoused and developed in seminal work by Nobel Laureate Elinor Ostrom [12]. Yet when we interpret this theory and try to put it into practice there are many unanswered questions. What makes the best criteria for peer group composition ... geography, social networking, demographics, virtual ? Should peer group membership be anonymous or revealed ? How should

we manage the free riders, should we just tolerate or expose them ? Can or should people be members of more than one peer group ?

These issues are deep questions for any attempt to use peer groups and social norms to persuade people to change their behaviour. Meanwhile, in our own study we don't have access to 170,000 users but we did develop two interventions in our feasibility trial of domestic energy usage which exploited provincial norms, and two interventions which did not. In the next section we introduce these 4 interventions which are each based on current technology available today.

4 Four Feedback Interventions for Electricity Consumption

4.1 Intervention 1 – Simple Web Interface

From the onset, our households expressed a strong desire to have web access to their data. Therefore as a natural initial step to maintain household interest and momentum, a simple web chart was developed. Data was captured into a central database repository and displayed online as a web application.Households were able to select their (known but anonymous) ID number from a drop-down box to then view near real time electricity consumption (in KW/h units) via a web browser.Similar near real-time feedback of usage is now available with systems developed and marketed by companies like CurrentCost[1] and AlertMe[2]. While this intervention did have limited accessibility, it did help to build user trust in the overall logging system,

4.2 Intervention 2 - In-Home Touchscreen Display

Interaction platforms such as a touchscreen in-home display, wall-embedded screen or interactive digital picture frame all offer possibilities for providing continuous, always-on, *ambient* feedback of electricity usage. This satisfies many of the user interface design criteria mentioned earlier and thus a simple standalone application running on a touchscreen computer was our next choice of display platform as it allowed us to develop and deploy, then define and re-deploy. Figure 1 shows a screenshot from the application showing 'day view' (upper panel) with the hourly breakdown of today's usage in bars and grey line indicating the average usage of that time in the past, and 'hour view' (lower panel) showing the minute breakdown in thin bars. Bar charts provide historical (past) information along with the current real-time usage. The user can click on the arrow buttons (right end of the top bar chart) to change the current focus to other hours and the bar charts scales (default is 'day view') to weekly or monthly views. Cost information (measured in Euro) is provided providing baseline and contextual information in an easily appreciated form. Each horizontal bar in all the charts is colour-coded so that if the usage is below the user's average it is in yellow-green, and if it is above average it turns to orange.

[1] `http://www.currentcost/com`
[2] `http://www.alertme.com`

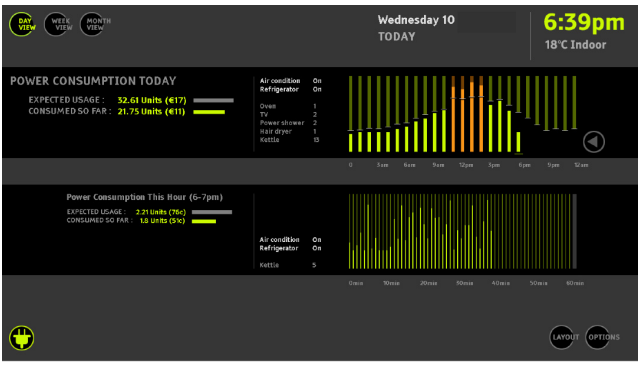


Fig. 1. The touchscreen display

4.3 Intervention 3 - Monthly e-Mail Awards

To build up peer groups, we devised a monthly e-mail awards scheme among the group of 24 households in our study who were divided into three social groups based on geographic region thus building on local rivalries among the groups. Individuals were rewarded for the savings made in comparison to their previous month, contributing to their group score. The e-mails informed individuals of their relative percentage savings and drew comparisons between themselves and their peers and their rivals in the other groups, and this was done in an encouraging and humorous manner.

4.4 Intervention 4 - Public Web Portal

In our final intervention we provided our households with anytime anywhere real time access and allow households to review how they were doing in comparison to everyone else. Figure 2 shows the Silverlight web interface we developed. Initially the user is presented with her own hourly electricity usage as a bar chart, with yellow-green bars indicating under-use and orange bars indicating over-use compared to her own norms. The user can change the visualisation to daily, monthly, or yearly and can also look at the other 23 anonymised households by selecting from a drop-down box thus allowing comparison with others. Household location is indicated on a map on the right panel, providing a sense of proximity with other households being compared against.

5 Experimental Setup

To monitor electricity consumption of each individual household, we make use of the commercially available EpiSensor ZEM-30[3] data logging unit illustrated

[3] `http://www.episensor.com/products/wireless-nodes/zem-30/`

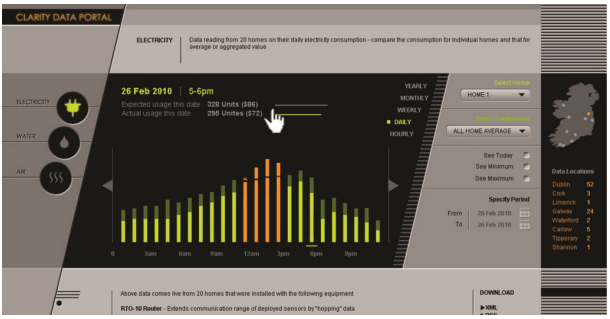


Fig. 2. Sensor Data Portal

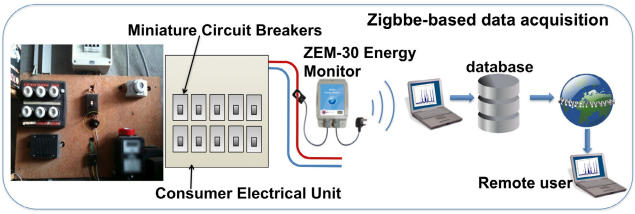


Fig. 3. Overview of capturing home electricity usage data and uploading to a central server

in Figure 3. The ZEM-30 includes a mains power supply, waterproof enclosure and a CT clip which is attached around the live wire running into a given household's main fusebox. This measures 11 different electrical parameters including *RMS/Peak/SAG current/voltage, real/apparent power* and of most relevance *Watt hours*. This information is relayed across a local Zigbee network to a computer every minute which logs the data onto a local relational database. This computer then uploads data every 10 minutes to a central web server, thus backing-up the data and also making it available for access from anywhere.

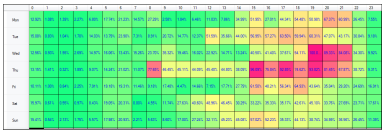
For a trial of 18 months duration, we recruited 24 households to log minute-by-minute electricity consumption data. The households differ in terms of geography (city vs rural), demographics (families, couples and singles), and household type (house, apartment). The captured data amounted to over 67.9 million parameter readings. Tables 1 and 2 summarise the data collected, with the "Typical Weekly Consumption" column visually shows the widely varying electricity consumption lifestyle patterns across all the households[4], with the x-axis being the 24 hours in a day and the y-axis being the 7 days in the week (Monday to Sunday, top to bottom). Households with a * in the table indicate those used to evaluate methods mentioned later in the paper.

[4] Household 3 is in fact a University chemistry lab, hence the concentration of electricity usage between 9am and 5pm, Monday to Friday.

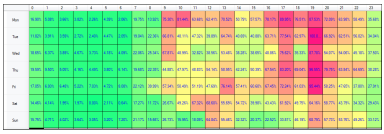
Table 1. Electricity data captured by 18 of our 24 households

Home # (number of occupants) and Visualisation of Typical Weekly Consumption

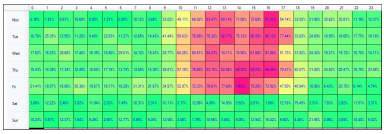
1(4)



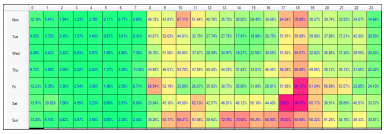
2*(2)



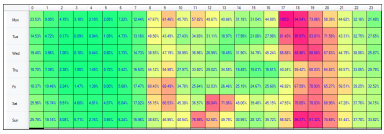
3*(10)



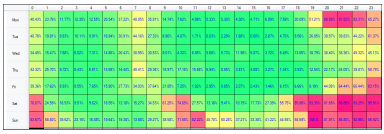
4*(3)



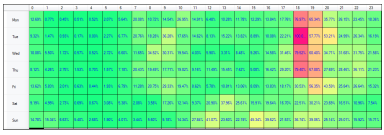
5*(4)



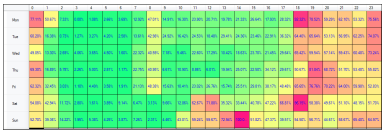
6*(1)



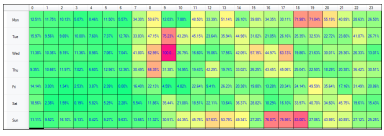
7(3)



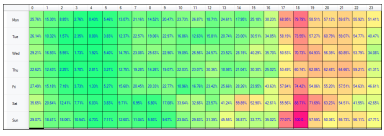
8*(2)



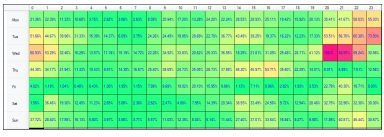
9(3)



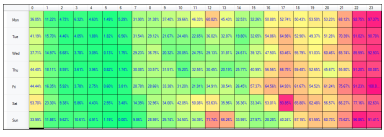
10*(5)



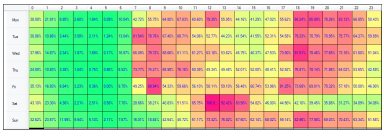
11(2)



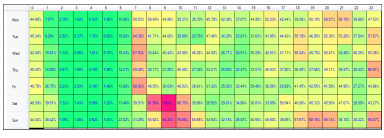
12*(2)



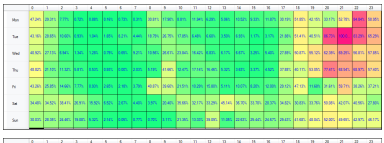
13*(4)



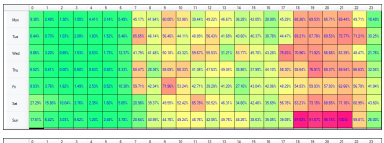
14*(3)



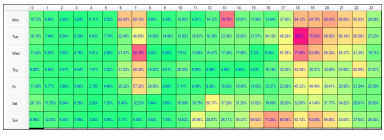
15*(3)



16*(5)



17*(2)



18(4)

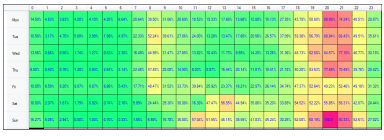
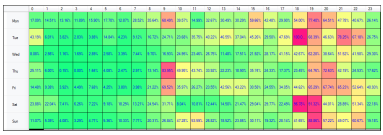
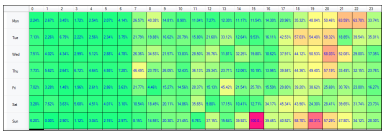
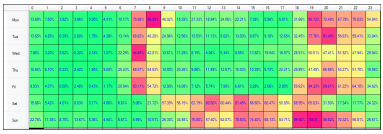


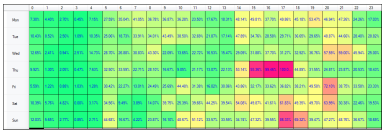
Table 2. Electricity data captured by remainder of our 24 households

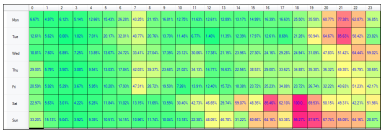
Home # (number of occupants) and Visualisation of Typical Weekly Consumption

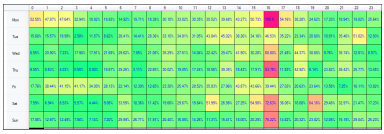
19(3) 

20(5) 

21(1) 

22(5) 

23(2) 

24(5) 

6 Results of Feasibility Tests

6.1 Intervention 1 Anecdotes - Simple Web Chart

As this stage was just an interim measure to help learn typical consumption behaviours, we did not gather data from households. However it is worthwhile to note that a handful of the households (only 3 or 4) accessed the web page multiple times per day, however the others infrequently accessed the data.

6.2 Intervention 2 Anecdotes - Home Computer Display

When many of our households were initially recruited, we captured some background usage information, so as to build up a profile of their energy consumption behaviour. Thereafter, we presented our users with the touchscreen application and noted savings in the months afterwards. Our 13 regular households (marked with a * in Tables 1 and 2) saved 14% on average one month after receiving access to the desktop display, 9% two months later, 14% three months later, and 16% four months later.

6.3 Intervention 3 Anecdotes - Monthly Email Awards

Monthly e-mails with awards for most saving were administrated for the three different social groupings. These generated email discussion but did not generate any noticeable effect on electricity consumption.

6.4 Intervention 4 Anecdotes - Public Web Portal

When our web portal was deployed, all our households were informed of its availability and encouraged to use it. There was no measurable effect on the

electricity consumption habits of our 13 most regular households, with an average daily consumption for the 2 months before the introduction of the web portal of 11.58 units vs. 11.23 units for the 2 months after this intervention was introduced.

7 Discussions

Touchscreen Display (Method 2): Anecdotally all households regarded our desktop-display as being easy to use. Households did consume less electricity after receiving access to the desktop display, however these savings could be as much to do with seasonal effects as access to the information. We could find no strong evidence that more usage of the browser has a direct effect on electricity consumption. Compared to our **initial web page (method 1)**, the computer display supports more continuous, less-effortful access to electricity monitoring, but even following best practice HCI guidelines, there still remains the significant challenge in supporting households to sustainably break the 5-15% savings barrier.

Monthly e-Mails Awards (Method 3): The primary challenge was to maintain sensitivity towards those households who saved the least in a given month, i.e. good performances were highlighted and praised while bad performances were not commented on. Initially there was much fun with many group e-mails being sent around and well-spirited comments showing competitive instincts. However, after initial interest, our households rapidly became less engaged over time. Winning households in the later months still felt a strong sense of achievement but as with the touchscreen display, the monthly e-mail awards failed in breaking the sustainable 5-15% savings barrier.

Web Portal (Method 4): The biggest challenge with this intervention has been in generating excitement among the households and getting them to use it. Our households have anecdotally noted that they use it to show to their friends but our 13 regular households appear to have hit a barrier in terms of further reducing their consumption. Perhaps an easier access through a mobile app may have helped here but the energy savings (just 3%) have plateaued during usage of this intervention.

8 Conclusions

Much literature has focused on the savings that can be made via continuous feedback of electricity consumption, but there is a lack of research focusing on the details of what those feedback mechanisms should consist of. In this article we concentrated on the design and implementation of four feedback mechanisms, discussing the strengths and weakness of each on a pilot sample of households. Two of these were based on exploiting peer norms that come from membership of peer groups and two were independent. The interventions have not been evaluated as behaviour change interventions due to a number of factors including the absence of control/baseline data, no structured randomised control trial to

assign deployments to participants, the elimination of almost 45% of the original study cohort from data analysis due to various technical issues, no adjustment for confounding factors such as weather, season, etc., and the limitations in deploying four interventions in succession without controlling for learning effects and other interactions between conditions that might confound quantitative results. So with all those negatives, is there anything we have learned from this work?

Through this experience we have found that our proposed interventions are feasible to deploy. Furthermore no one intervention method was preferred by all the participants, rather a combination of the methods allowed the participants access their electricity consumption information depending on the context of the circumstances around them. We have also found that a variety of feedback mechanisms and interventions are needed in order to sustain user interest and that the challenges of changing user behaviour in this domain still remain. While peer grouping which exploits "Big Data" from smart metering is recently shown to be scalable [1], the composition of groups, membership visibility, ability for users to opt in/out of groups etc., all need to be addressed. However, the rewards of getting this right will be significant in terms of reduced carbon emissions and reduced consumer energy bills.

Acknowledgements. The authors thank Science Foundation Ireland (grant 07/CE/I147) and (AD) the Irish Health Research Board (grant MCPD/2010/12).

References

1. Ayres, I., Raseman, S., Shih, A.: Evidence from two large field experiments that peer comparison feedback can reduce residential energy usage. Journal of Law, Economics, and Organization (2012) (Online: published August 20, 2012)
2. Chetty, M., Tran, D., Grinter, R.E.: Getting to green: understanding resource consumption in the home. In: UbiComp 2008: Proceedings of the 10th International Conference on Ubiquitous Computing, pp. 242–251. ACM, New York (2008)
3. Dada, A., Staake, T., Fleisch, E.: The potential of UbiComp technologies to determine the carbon footprints of products. In: Workshop on Pervasive Persuasive Technology and Environmental Sustainability, Pervasive 2008, pp. 50–53 (2008)
4. Darby, S.: The effectiveness of feedback on energy consumption. Technical report, Environmental Change Institute, University of Oxford (2006)
5. Fischer, C.: Feedback on household electricity consumption: a tool for saving energy? Energy Efficiency 1(1), 79–104 (2008)
6. Fitzpatrick, G., Smith, G.: Technology-enabled feedback on domestic energy consumption: Articulating a set of design concerns. PERVASIVE Computing: Special Issue on Environmental Sustainability 8(1), 37–44 (2009)
7. Goldstein, N.J., Cialdini, R.B., Griskevicius, V.: A room with a viewpoint: Using social norms to motivate environmental conservation in hotels. Journal of Consumer Research 35(3), 472–482 (2008)
8. Hazas, M., Friday, A., Scott, J.: Look back before leaping forward: Four decades of domestic energy inquiry. IEEE Pervasive Computing 10(1), 13–19 (2011)

9. Holmes, T.G.: Eco-visualization: combining art and technology to reduce energy consumption. In: C&C 2007: Proceedings of the 6th ACM SIGCHI Conference on Creativity & Cognition, pp. 153–162. ACM, New York (2007)
10. Martin, G.: Smart metering information paper 4 - results of electricity cost-benefit analysis, customer behaviour trials and technology trials. Technical report, Irish Commission for Energy Regulation (2011)
11. McCalley, L.T., Midden, C.J.H.: Energy conservation through product-integrated feedback: The roles of goal-setting and social orientation. Journal of Economic Psychology 23(5), 589–603 (2002)
12. Ostrom, E.: Collective action and the evolution of social norms. The Journal of Economic Perspectives 14(3), 137–158 (2000)
13. Petersen, D., Steele, J., Wilkerson, J.: Wattbot: a residential electricity monitoring and feedback system. In: CHI 2009: Proceedings of the 27th International Conference Extended Abstracts on Human Factors in Computing Systems, pp. 2847–2852. ACM, New York (2009)
14. Petroni, P., Cotti, M., Bono, O.: The new edge for the enel telegestore: An integrated solution for the remote management of electricity and gas distribution allowing a total management of the energy consumptions. In: 2009 20th International Conference and Exhibition on Electricity Distribution, pp. 1–5 (June 2009)
15. Pierce, J., Schiano, D.J., Paulos, E.: Home, habits, and energy: examining domestic interactions and energy consumption. In: CHI 2010: Proceedings of the 28th International Conference on Human Factors in Computing Systems, pp. 1985–1994. ACM, New York (2010)
16. Riche, Y., Dodge, J., Metoyer, R.A.: Studying always-on electricity feedback in the home. In: CHI 2010: Proceedings of the 28th International Conference on Human Factors in Computing Systems, pp. 1995–1998. ACM, New York (2010)
17. Strengers, Y.: Smart metering demand management programs: challenging the comfort and cleanliness habitus of households. In: OZCHI 2008: Proceedings of the 20th Australasian Conference on Computer-Human Interaction, pp. 9–16. ACM, New York (2008)
18. The Olympic Peninsula Project. Gridwise demonstration project fast facts. Technical report, Pacific Northwest National Laboratory (2007)
19. Velte, D., Martin, A.: Energy services: Reducing the energy consumption of residents by behavioural changes. Technical report, BewareE (2009)
20. Weiss, M., Mattern, F., Graml, T., Staake, T., Fleisch, E.: Handy feedback: connecting smart meters with mobile phones. In: MUM 2009: Proceedings of the 8th International Conference on Mobile and Ubiquitous Multimedia, pp. 1–4. ACM, New York (2009)

Spotz: A Location-Based Approach to Self-awareness

Misha Sra and Chris Schmandt

MIT Media Lab., Cambridge, MA 02139

Abstract. This paper introduces the location-based mobile application Spotz that explores the persuasive qualities of sharing location information visually to promote behavior change. Spotz encourages users to become self-aware of the kinds of places they visit which can have motivational properties deriving from social feedback. The app displays a continually evolving graphic of relatively sized circles depicting the number and type of places at which the users check-in, including the option to upload this visual to social media.

1 Introduction

As the field of persuasive technology is focused on developing computing systems that attempt to change attitudes and behavior [1], employing self-reflection in persuasive technologies might be an effective means of supporting behavior change. The Spotz approach differs from previous location, exercise and social tracking applications by providing an opportunity for users to shape both how they see themselves and how they are perceived by habituating locations which are crowd designated as having a certain personality. Whereas previous applications use quantitative data gathering to create statistics of area covered or miles logged, Spotz builds a socio-location mirror of personal habits which users can examine and provide a pivot of self-persuasion for any number of changes from health to education and community involvement.

The importance of this work to Persuasive Technologies is twofold. First, this work presents a working smartphone location-based app that presents real world location check-in data to the user in ways that they can reflect upon and comprehend easily. Second, it presents the data in a way that makes it easily shareable across social media websites to build upon the complex idea of self-representation in online worlds for behavior change.

2 Previous Work

The abundance of smart mobile devices with sensors in an always-connected world enables the recording and sharing of a countless variety of personal data. If all this data are made available to users with an understanding of users' self-reflection needs [2] they can become more aware of their own behavior [3] and change behavior in areas ranging from health to sustainability [4,5]. People may

S. Berkovsky and J. Freyne (Eds.): PERSUASIVE 2013, LNCS 7822, pp. 216–221, 2013.

access and manage this data using various tools and make meaning of it with varying degrees of success for the purpose of self-knowledge [6]. For example. playful tools such as Fish'n'Steps [7] which ties progress in a virtual game to the number of steps walked by the user to Ubigreen [8] for affecting behavior change in the context of transportation, resource usage.

3 Personal Informatics

Personal informatics is a class of tools that help people collect personally relevant information for the purpose of self-reflection and self-monitoring. These tools help people gain self-knowledge about one's behaviors, habits, and thoughts [9]. A variety of data, both quantitative and qualitative about step count, physiology, moods or sleep are available to be collected and visualized. People may use this data for purposes of health management, self-motivation and record keeping.

There exists a lot of research on effective information visualization techniques to explore data for trends and patterns but not focused on personal data. There are a few projects that have applied visualization techniques on personal data. For example, Andry et al. have designed a diabetes visualization module for LifeSensor to optimize data collection as well as help patients explore relationships between their blood glucose levels and their diet [10]. Commercial products, such as Nike+, Fitbit or Motoactv (physical activity), Moodpanda (moods), Zeo (sleep), Klout (online influence), and Wattson (electricity), use information visualizations to assist users in exploring their data for information, patterns and self-motivation.

4 Application Design

Spotz is an easy to use mobile phone application for Android phones. It uses the familiar paradigm of 'checking-in' and builds visualizations for the user based on qualitative information about the places the user checks in at.

4.1 User Scenario

Jane is a college student in her mid twenties. She believes that she is a healthy eater and only goes to healthy restaurants and grocery stores who likes 'geeky' things. She downloads Spotz out of curiosity and every time she checks-in to Foursquare she does a parallel check-in to Spotz and reviews the results at the end of the week that surprise her.

Her user profile shows a visualization of the types of places she checked-in at during the week, a summary of the top places she visited, and recommendations for new places based on her personality as created by her use of the app. The visualization shows her as a predominantly geeky person with a little interest in healthy places contrary to her expectations. Reviewing the list of places she realizes that she has been going to the local cafe a lot more than she thought

she did. She recognizes that she rarely goes to healthy restaurants but instead favors the fast food restaurants at the nearby food court out of convenience. She wants to share her persona visualization with her friends on social media sites but decides to wait until she can alter the visual representation of the locations she visits by going to places that are representative of the type of places she wants herself to be associated with.

The app also allows her to 'follow' broad pre-programmed personality tags for places (healthy, fun, geeky) and user generated 'buzz' tags (best vegan chocolate cake, safe running trail) that are for very specific items or characteristics of a place. For e.g. following the 'buzz' trail of the vegan chocolate cake she may discover a neighborhood gas station store which would not be a normally recommended place for 'healthy' eating.

4.2 App Design

To ensure an easy to use app interface, the check-in part the app is modeled (Figure 1) after the popular location-sharing social app Foursquare. The intention was to create an app that is fun and easy to use yet is not obvious about its persuasive goals. Facebook was selected as the primary sharing platform given its popularity and an accessible API. As conventional understandings of human identity, representation, and social relations are being revised in the light of technological mediation [11], users now have means and reasons to share different facets of themselves, projecting a public self that may very well be different from the private self. By providing interesting information to the user about their check-in behavior in a playful manner, the app trusts self-reflection to emerge as the primary behavior change mechanic.

4.3 App Functionality

Users can check-in to places using the GPS hardware in the mobile device to display a list of nearby locations or by scanning NFC stickers at places. Check-in is an active user selection process following which the user is presented with the option of 'tagging' that location (Figure 2). It is this 'tagging' characteristic that forms the basis of building a location's personality and based on the location's personality, the personas of users who visit that location emerges.

4.4 Persuasive Elements

In Spotz, multiple persuasive strategies [12] have been implemented in various elements. The continually evolving user persona visualization was used as a new type of check-in data representation, different from the lists or representations on a map. The user participates in the creation of their persona in two ways. First, by visiting a location and 'tagging' that location the user assigns a personality to the place (Figure 2). Based on the predominant 'tag' for a location, its personality emerges, even though it's just a matter of time when people start

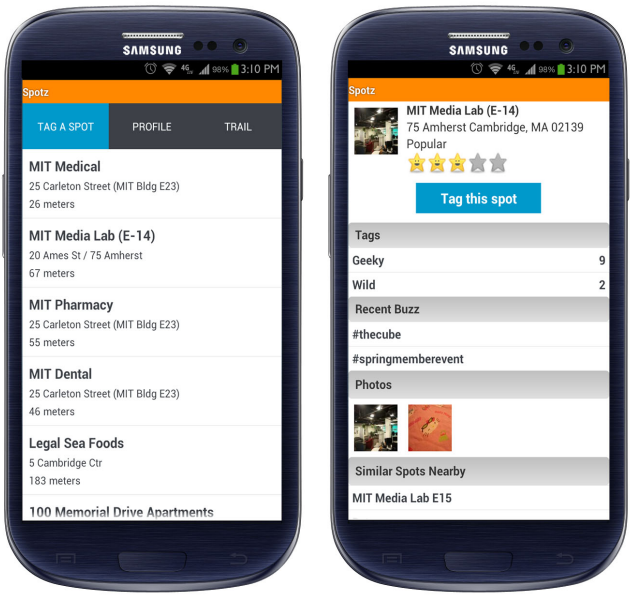


Fig. 1. User selects and tags their current location, contributing to creating a personality of the location which in turn affects the persona of the people who check-in at this location

'tagging' it differently and entirely change the places' personality. Second, when a user visits a place, they earn a 'tag', which is the same 'tag' as the location personality, thereby creating the idea of 'you are where you go.' In order to control the visualization of their persona, the user need only check-in at places with the personalities they want associated with them. For example, if a user wants to project a 'healthy' persona, they may choose only to check-in at places that are crowd tagged 'healthy', thereby earning a 'healthy' tag and thus increasing the size of their 'healthy' circle in the persona visualization or they could influence and change the 'nice' tag of their favorite place to 'healthy' by checking-in there frequently and asking friends to do the same.

In order to reward the user for making an effort to visit new places, the adventurous rating (Figure 2) was implemented as a new type of positive personality characteristic that people would want associated with them. The idea was to use a simple star rating mechanism that can at a glance provide information to the user thereby motivating them to visit newer places to increase their adventurous rating.

The following design strategies were incorporated: 1) simple paradigm of checking-in (familiarity); 2) adventurous rating (player empowerment); 3) the opportunity to 'tag' locations (player empowerment); 4) the opportunity to create special 'tags' for (Figure 3) characteristics of the location (player empowerment); 5) creating of user persona (Figure 3) based on their location check-in information (game interaction); 6) animated visual representation (Figure 3) of

Fig. 2. Left: Visualization shareable to Facebook and list of recent places tagged by user. Right: Location personality graph and 'buzz' associated with that location.

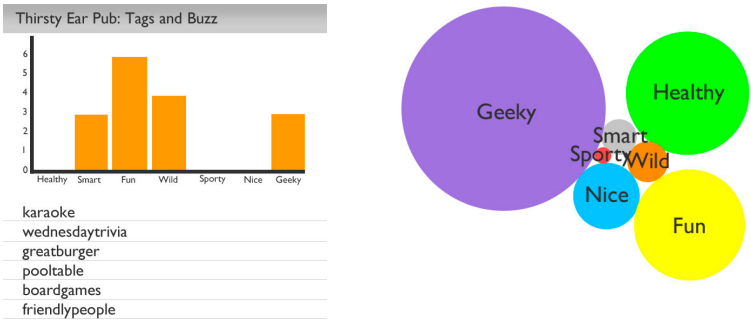


Fig. 3. Left: Enlarged location personality graph with 'buzz' associated with that location. Right: User persona visualization detail.

user persona (game interaction); 7) ability to 'follow' and receive information about new places nearby based on the location personality (timely information, minimum effort, low cognitive load); 8) the ability to 'follow' and receive information about specific objects (newness, exploration); 9) opportunity to share persona information online (social) (Figure 3).

5 Preliminary Evaluation

Spotz was given to a small group of student collegues and received positive verbal feedback on the design of the visualization. It was suggested to include more social elements specifically a friend feature which would require more attention to privacy and security of user data. However, to move the project foward, a collaboration with existing location-based apps would be ideal as this work does not aim to reinvent what has already been successfully done but to provide a different perspective on viewing location information and providing the opportunity to reflect upon data.

References

1. Fogg, B.J.: Persuasive Technology: Using Computers to Change What We Think and Do. Morgan Kaufmann Publishers, San Francisco (2003)
2. Li, I., Dey, A.K., Forlizzi, J.: Understanding My Data, Myself: Supporting Self-Reflection with Ubicomp Technologies. In: UbiComp 2011 (2011)
3. Carver, C., Scheier, M.F.: On the self-regulation of behavior. Cambridge University Press (2001)
4. DiClemente, C.C., Marinilli, A.S., Singh, B., Bellino, E.: The Role of Feedback in the Process of Health Behavior Change. American Journal of Health Behavior 25(3), 217–227 (2000)
5. Seligman, C., Delay, J.M.: Feedback as a Means of Decreasing Residential Energy Consumption. Journal of Applied Psychology 62(4), 363–368 (1977)
6. Li, I., Dey, A.K., Forlizzi, J.: A Stage-Based Model of Personal Informatics Systems. In: CHI 2010, pp. 557–566 (2010)
7. 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)
8. 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 (2009)
9. `http://personalinformatics.org/`
10. Andry, F., Naval, G., Nicholson, D., Lee, M., Kosoy, I., Puzankov, L.: Data Visualization in a Personal Health Record using Rich Internet Application Graphic Components. In: HEALTHINF 2009, pp. 111–116 (2009)
11. Yaszek, L.: The self wired: technology and subjectivity in contemporary narrative. Routledge, New York (2002)
12. Khaled, R., Biddle, R., Noble, J., Barr, P., Fischer, R.: Persuasive Interaction for Collectivist Cultures. In: Proc. of The Seventh Australasian User Interface Conference, vol. 50, pp. 73–80 (2006)

Bro-cam: Improving Game Experience with Empathic Feedback Using Posture Tracking

Chiew Seng Sean Tan, Johannes Schöning, Jan Schneider Barnes,
Kris Luyten, and Karin Coninx

Hasselt University
tUL – iMinds
Expertise Centre for Digital Media
Diepenbeek, Belgium
firstname.lastname@uhasselt.be

Abstract. In todays videogames user feedback is often provided through raw statistics and scoreboards. We envision that incorporating empathic feedback matching the player's current mood will improve the overall gaming experience. In this paper we present Bro-cam, a novel system that provides empathic feedback to the player based on their body postures. Different body postures of the players are used as an indicator for their openness. From their level of openness, Bro-cam profiles the players into different personality types ranging from introvert to extrovert. Empathic feedback is then automatically generated and matched to their preferences for certain humoristic feedback statements. We use a depth camera to track the player's body postures and movements during the game and analyze these to provide customized feedback. We conducted a user study involving 32 players to investigate their subjective assessment on the empathic game feedback. Semi-structured interviews reveal that participants were positive about the empathic feedback and Bro-cam significantly improves their game experience.

1 Introduction

The recent push towards using gestural and motion gaming platforms (such as Microsoft Kinect, Sony PlayStation Move, and Nintendo Wii) uses the player's physical actions for the interaction with the game. Still, game feedback is presented in raw numbers and is not adapted to the user's mood. We expect that adapting the feedback to the current mood of the players will enhance the overall game experience as users have shown the preference for an empathetic agent with non-repetitive dialogs [3]. For example, it would be more appropriate to support users with encouraging comments after a game when they are slouching their shoulders instead of making harsh cynical comments. This is also motivated by recent work of Aviezer et al. [1]. They have shown that body postures are typically used more often than other modalities for recognizing emotions. In their Science paper Aviezer et al. analyzed body cues and facial expressions of different professional tennis players. Their study showed that

S. Berkovsky and J. Freyne (Eds.): PERSUASIVE 2013, LNCS 7822, pp. 222–233, 2013.

body cues, which dominate over facial expressions, are useful to discriminate between intense positive and negative emotions.

In this paper, we present our work on Bro-cam, a system that couples personalization and persuasion to addresses this challenge. Instead of using affective state to adjust gameplay, we describe a game "companion" system that observes the gameplay and the player's postures and then provides humorous feedback. Bro-cam provides its feedback to a player after each match (several matches are part of one gaming session). Players use their body postures and actions to interact with the game. Even when a match of a gaming session is over, the body posture can provide behavioral information about the player. This posture is used by Bro-cam as an indicator of the openness to present adequate feedback. Openness refers to one's personal tendency to take into consideration the possibility of accepting different (or novel) ideas and information, which may lead to a change in one's behavior accordingly. Our system tries to "understand" the player's mood after the game and tries to influence their attitudes by providing encouraging feedback. This is done by taking into account the players' level of openness and profiling them into different personality types. In a prestudy we investigated how rather extrovert or rather introvert players respond to different types of humorous feedback. Subsequently, Bro-cam personalizes the feedback by matching the personality type to their preferences for certain humoristic statements with the aim of increasing the players' motivation for the game.

The two main contributions of this paper are: (1) A system providing empathic game feedback, using the body posture of the players as an indicator of their openness to receive feedback. This approach is unique and novel, because we do not need any other type of user instrumentation, e.g. physiological sensors, to gather information about their level of openness. Furthermore, we can use the same hardware that is also used to interact with the game. (2) A user evaluation demonstrating that adapting game feedback to the players' openness is well received and enriches the overall game experience and the players motivation.

2 Related Work

Affective computing has been a widely investigated topic for years in human-computer interaction. The main research challenges and criticisms of affective computing are summarized by Picard [17]. In her cognitive models of emotion [16], the emotions are the central components in the system. In contrast to Picard, Höök and Gaver [11, 9] concentrate on using emotions as just one component contributing to the overall design goal of the application or system as also shown by Kuikkaniemi and Nacke [12, 14]. Many researchers have shown the huge potential of including affective states of users in human-computer interaction within various domains (e.g. games, education, security scenarios) [13].

Using physiological data to derive affective states as game input is not entirely new. Recently Nacke [14] demonstrated direct and indirect physiological sensor input

to enhance traditional game controls similar to the work of Kuikkaniemi et al. [12]. In addition, Berkovsky [2] developed *PLAY,MATE!* game design principles which take in account the user's existing engagement to seamlessly motivate them to perform physical activity while playing the game. Earlier works by Dekker [6] and Gilleade [10] provide good overviews of the research area of affective games.

Similar to our work, Gamerini [8] also tailor the game feedback by taking into account the users' personal (energy consumption) behavior. Their work differs from ours by using users' actions to trigger contexualized feedback in order for them to learn about energy conversation. Mapping of personality to the user's behavioral choice has been studied by Tapus et al. [20]. They reported that participants with extroverted personalities had a preference for a robot that challenged them during the training program. Whereas users with introvert personalities preferred the robot that focuses on nurturing praises. In contrast to related work our system passively observes the users' postures (openness) and personalizes the game feedback for motivating them to continue playing the game. We present a new unobtrusive technique to detect the affective states of the player. The same hardware, in our case the Microsoft Kinect, is used to track the gesture input during the game and to give feedback that adapts to the affective state of the users after the game.

Following the abovementioned arguments of Höök and others, we use the affective state of users as input to provide empathic feedback after the gameplay. We see huge potentials in enhancing the game experience with feedback adapted to users' affective state using the same hardware that is already used in motion- or gesture-based controlled games.

3 Bro-cam System Overview

The Bro-cam system consists of three main components, which we describe in the following subsections. The first component is the posture classifier that infers the players' openness level. The second component maps the game outcome and player's behavior to appropriate feedback type. This is supported by a pre-study, in which we investigated how rather extrovert or rather introvert players respond to different humoristic statements in winning or losing situations. The feedback statements are then generated in the third component and presented to the player on an additional screen.

3.1 Inferring Openness from Postures

We used a vision-based posture recognition approach to track users' body joints using the Microsoft Kinect depth camera. We implemented a combination of temporal scaling and spatial transformation parameters for 3D body modeling. A 3D body model is tailored for detecting postures based on the OpenNI SDK from PrimeSense[1].

[1] www.openni.org

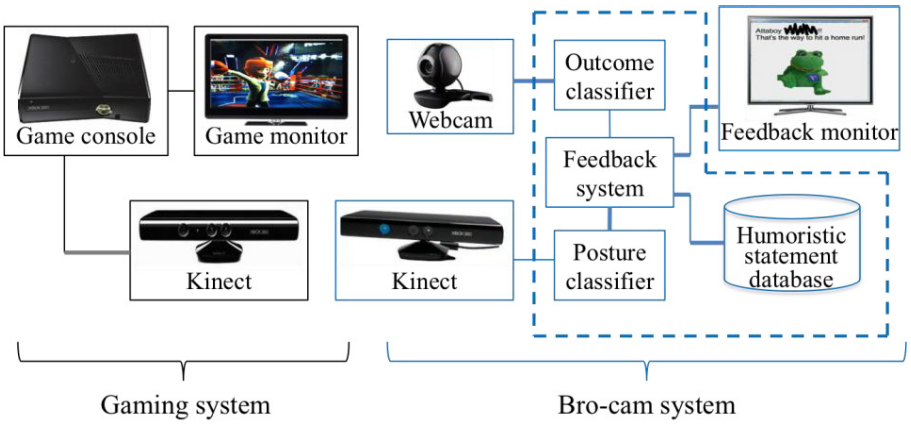


Fig. 1. Bro-cam setup. Two Kinects (stack on top of each other) are facing the player. A web-cam is pointing directly at the game monitor to determine for the win-lose outcome. The feed-back statements, which are personalized to the player's behavior and win-lose outcome, are then presented to the player on an additional screen.

We extracted the angles of fourteen monitored body joints and tracked the move-ment of these joints every 200 milliseconds. A feed-forward neural network with a Backpropagation Learning algorithm executes the posture recognition algorithms (one for each defined posture).

With this posture classifier approach we are able to successfully classify eleven postures with a mean accuracy of 84.98% (standard deviation of 7.41%) and RMSE of 1.89%. These eleven postures – typically identified in nonverbal communication – are (1) Hand-on-chin, (2) Hand-near-hip, (3) Hand-behind-neck, (4) Arm Crossed, (5) Limbs Restrained, (6) Limbs Relaxed, (7) Legs Crossed, (8) Foot Forward (Left), (9) Foot Forward (Right), (10) Lean Forward, and (11) Lean Sideward.

Using a ten-fold cross validation on the training data, the classifier models yield average accuracy rate of 91.03% for Hand-on-chin, 87.85% for Hand-near-hip, 85.97% for Hand-behind-neck, 81.38% for Arms Crossed, 95.23% for Limbs Re-strained, 83.44% for Limbs Relaxed, 78.56% for Legs Crossed, 74.9% for Foot For-ward (Left), 96.06% for Foot Forward (Right), 86.14% for Lean Forward, and 74.17% for Lean Sideward. More details can be found in [19].

To confirm the level of openness conveyed by each of the postures we asked human observers to assess their emotional content. We manually annotated various examples of each of the postures as test patterns and 60 observers (recruited from Amazons Mechanical Turk) rated each test pattern on its level of openness from 1 (not approachable at all) to 7 (very approachable). The faces in the test patterns were pixelated to prevent observers from interpreting emotion from facial expression. The interrater reliability using the Cohen's Kappa statistic, Kappa = 0.59 ($p < .001$), 95% CI (0.468, 0.712), indicated moderate strength on the agreement among the observers.

Table 1. Profiling of players according to openness level based on human observers rating on 7-point Likert scale. Techniques for character interpretation from Blumenfeld [4] and affect evaluation from Eysenck et al. [7] are used to guide our mapping of openness levels to personality types. Note that we assumed that the interacting entity is located at the direction where the foot forward is point to the right side i.e. Foot Forward (Right).

Observers rating (7-point Likert scale)	Openness level	Personality type	Posture classification
$M = 4.43$, $S.D. = 1.54$	High	Extrovert	Hand-near-hip, Limbs Relaxed, Foot Forward (Right), Lean Forward
$M = 3.85$, $S.D. = 1.57$	Mid	Ambivert	Hand-behind-neck, Limbs Restrained, Lean Sideward
$M = 3.36$, $S.D. = 1.49$	Low	Introvert	Hand-on-chin, Arms Crossed, Legs Crossed, Foot Forward (Left)

The between-groups multivariate ANOVA results show that the postures can be grouped into three distinct levels. Postures with high level of openness (M = 4.43, SD = 1.54) were indeed rated significantly higher than postures with mid level (M = 3.85, SD = 1.57), $F(59) = 0.476$, $p < .001$; $r = 0.002$. Similarly, postures with low level of openness (M = 3.36, SD = 1.49) were rated significantly lower than postures with mid level, $F(59) = 0.467$, $p < .001$; $r = -0.113$. We then related the high level of openness to extroverts with positive affect evaluation [7] who appear approachable, friendly and sociable based on the character interpretation techniques from Blumenfeld [4]. Conversely, a low level of openness is associated with introverts with negative affect evaluation who also appear withdrawn, shy, and reserved, sometimes to the point of inscrutability. We refer to participants who exhibit both extrovert and introvert as ambiverts and associate them with a mid level of openness. Table 1 presents the profiling of players based on the observers rating into different personality type.

3.2 Feedback Mapping

To investigate how extrovert or introvert players respond to different humoristic statements in different situations we conducted a pre-study. We surveyed pre-test participants to find out their preference through subjective ranking on the appropriate feedback type in winning or losing situations. We asked 65 pre-test participants (recruited from the public fitness center, the cafeteria and the lobby of our university) to fill out the Eysenck Personality Questionnaire [7]. In addition we asked them to rank a compilation of feedback statements for winning or losing situations.

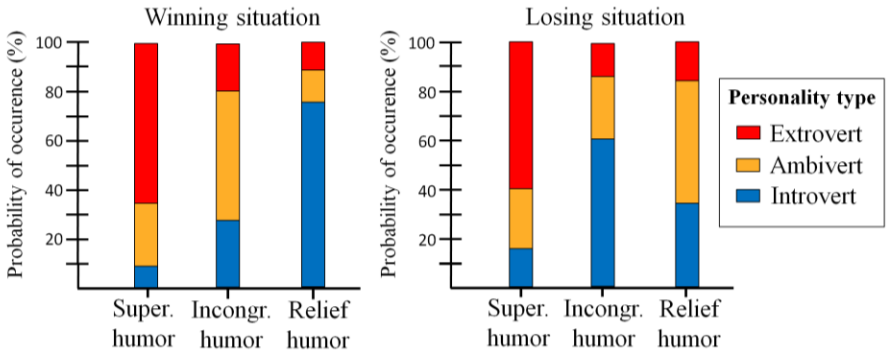


Fig. 2. The preference for each humor during winning and losing situation is presented in term of probability of occurrence as indicated by participants exhibiting one of the three personality types (extrovert, ambivert, introvert).

These humoristic statements are based on three main theories of humor (Superiority, Incongruity, and Relief), which have emerged primarily from psychological studies and research in emotion-oriented systems [18]. The superiority theory suggests that humor is a form of expressing the superiority of one person over another and laughter is triggered by our feelings of superiority with respect to others, e.g. "Atta-boy [name]! That's the way to hit a home run!". The incongruity theory suggests that humor is due to the mixing of two disparate interpretation frames for one event e.g. "You must feel as good and refreshed as an egg in the microwave after this beating". The relief theory suggests that humor is a form of bypassing certain censors that prevent us from having uncomfortable thoughts. Thus laughter is induced as a result of release of physical energy, which is built up to deal with disagreeable feelings, e.g. "You buy me a drink and I'll ask your opponent to let you win the next round". We used the findings from the survey to discover the type of humorous feedback that an extrovert (more open) or introvert (less open) would like.

The results on the pre-study, which indicates the preferred humor according to the three personality types, are illustrated in figure 2. We perform the Chi-Square test for independence to determine whether the personality type (extrovert, ambivert, introvert) is actually associated with the preferred humor (Superiority, Incongruity, Relief). In the winning situation, we can conclude from the Chi-Square test results that there is a statistically significant difference between the three personality types and the preferred type of humor. The results for this test are as follows: (1) Superiority humor versus Incongruity humor, $\chi^2(4, N=65) = 73.753$, $p = .001$, Phi = .859, V = .607, (2) Superiority humor versus Relief humor, $\chi^2(4, N=65) = 17.763$, $p = .001$, Phi = .421, V = .298, and (3) Incongruity humor versus Relief humor, $\chi^2(4, N=65) = 81.476$, $p = .001$, Phi = .903, V = .638. The same conclusion to the winning situation can also be drawn for the losing situation. The Chi Square test results for losing situation are as follow: (1) Superiority humor versus Incongruity humor, $\chi^2(4, N=65) = 44.429$, $p = .001$, Phi = .667, V = .471, (2) Superiority humor versus Relief humor, $\chi^2(4, N=65) = 79.184$, $p = .001$, Phi = .890, V = .629, (3) Incongruity humor versus Relief humor, $\chi^2(4, N=65) = 111.947$, $p = .001$, Phi = 1.086, V = .768 .

Table 2. Findings of humor type to the game outcome (winning versus losing) based on exhibited behavior of the player

	Extravert behavior	Ambivert behavior	Introvert behavior
Winning situation	**Superiority humor**	**Incongruity humor**	**Relief humor**
Losing situation	**Superiority humor**	**Relief humor**	**Incongruity humor**

Phi and Cramer's V are used as tests of the strength of association. We can see the average Phi = .804 and average Cramer's V = .569 indicate that the strength of association between the variables is fairly strong. We then create a mapping based on the findings of this pre-study, which is shown in table 2.

3.3 Feedback Presentation

We developed a feedback system that dynamically changes its output based on the player's extrovert-introvert behavior as indexed by the openness level and whether they won or lost the game. Once the behavior is obtained we need to know the game outcome. We developed an outcome classifier to automatically determine the game outcome by visually processing the game screenshots using a webcam facing the game monitor. This is simply done by analyzing the color histogram of the game monitor as can be seen in figure 1. The outcome classifier is pre-configured for recognizing the winning and losing color histogram, which in turn, is used to detect the game outcome (i.e. winning or losing status).

Following, we combined the openness level with the game outcome every 3 seconds and added an expiry function to allow extrovert-introvert behavior information to change over a moving window of 30 seconds. The expiry function is intended to reflect the player's dynamic behavioral transition during the progression of the game. With a decreasing marginal utility curve we define the gradual loss in the player's behavioral information. The most appropriate humoristic statement is then selected based on total match strength of the player's behavioral information with the game outcome to create personalized feedback in the form of a sentence. Consequently, the sentence is presented to the player via an avatar (as shown in figure 1). This avatar serves as a social buddy to establish a friendly relationship with the player.

4 Evaluation

To evaluate the Bro-cam system for the appropriateness of the feedback we refer to the user satisfaction rather than what the system logs. A user study was conducted with (subjective and objective) quantitative measures. We recruited 36 participants, who are used to play video games. The study was successfully completed by 32

participants (8 female, 24 male). Data from 4 participants has been discarded, as they did not complete the game due either to personal or health issues. The participants were students or staff at our university and have varying degrees of educational background. Their ages range from 18 to 33.

4.1 Apparatus

The experiment setup is shown in figure 1. The Microsoft Xbox360 gaming console was connected to a 32-inch full HD display. The feedback of the Bro-Cam system was provided on a 42-inch full HD display. Both Kinect sensors were mounted directly below the center of the game monitor. The participants stood about six to eight feet away from the Kinect sensor. The study took approximately 25 minutes for each session. The participants played two matches of Microsoft "Kinect Sports" videogame "Boxing" at the beginner level. The introduction video for the "Kinect Sports" videogame was shown to all participants before they played the game.

4.2 Procedure

In the first match of the game, we measured the time to complete the match and established a performance baseline. We then allowed the participants to rest until they felt comfortable to continue with the second match. In the second match of game, we measured again the time taken to complete the match and compared it with the performance baseline to determine the skill level. After each match a feedback from the Bro-cam system was presented to the players. At the end of the session, we conducted a semi-structured interview with the participants to assess qualitatively the impact of empathic feedback. The participants were then debriefed and were allowed to ask any additional questions.

5 Results and Analysis

We perform a set of quantitative analysis on the participants' subjective and objective response, which is obtained from their post interview and game performance respectively. The participants' Likert scale ratings are examined to determine how well the empathic feedback matches the game experience and also how well this feedback motivates them to continue playing with the game.

We assess the strength of agreement from the participants' subjective response using Cronbach's α. The result indicates a fairly high strength of agreement score ($\alpha = 0.668$) on the internal consistency for the set of data from the post interview. We then compute for the Levene's Test for Equality of Variances which shows that the variability of the three personality types (extrovert, ambivert, introvert) is the same for the two quantitative measures: (I) feedback is a match to game experience; and (II) feedback is motivating for gameplay.

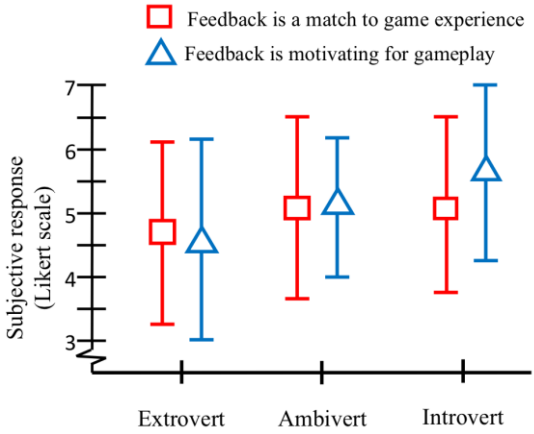


Fig. 3. Likert scale ratings for feedback from players exhibiting different personality type. A Likert scale of 1 to 7 (higher is better) is used. Error bars show standard deviation.

The ANOVA test for (I) shows $F(29) = .003$, $p = .960$, which we can conclude that there is no significant difference in term of how the empathic feedback match their game experience between extrovert ($M = 4.731$, $SD = 1.485$), ambivert ($M = 5.091$, $SD = 1.509$), and introvert ($M = 5.062$, $SD = 1.436$) participants. Similarly, the ANOVA test for (II) shows $F(29) = 1.494$, $p = .229$, which provide us with the statistical evidence that there is also no significant difference in term of how their motivation for gameplay is affected by the empathic feedback between extrovert ($M = 4.654$, $SD = 1.623$), ambivert ($M = 5.182$, $SD = 1.097$), and introvert ($M = 5.687$, $SD = 1.352$) participants.

We used Wilcoxon signed rank test using median equals 4.0 to determine whether there is significant matching of the empathic feedback to the participant's game experience. The extrovert participants with $p = .0095$ has shown a significant effect ($p < .05$) on matching the empathic feedback to their game experience. The ambivert participants ($p = .0015$) and introvert participants ($p = .005$) has shown similar significant effect on matching the feedback to their experience.

Likewise, the results of Wilcoxon signed rank test using median equals 4.0 also show that there is significant improvement on their gaming experience from the empathic feedback. The extrovert participants with $p = .025$ has shown a significant effect ($p < .05$) on matching the empathic feedback to their game experience. The ambivert participants ($p = .001$) and introvert participants ($p = .001$) has shown similar significant effect on matching the feedback to their experience (figure 3).

To analyze the objective quantitative measure from the gameplay we focus on the player's skill level, which according to Pfeifer [15], varies the game experience of the players accordingly. After a full analysis on all the players' score and timing, we can categorize 32 participants into two groups which varies according to their skill level: above-average players and below-average players. The skill levels are based on their final scores on each match. Participants with high scores are identified as above-average players and those with lower scores as the below-average players.

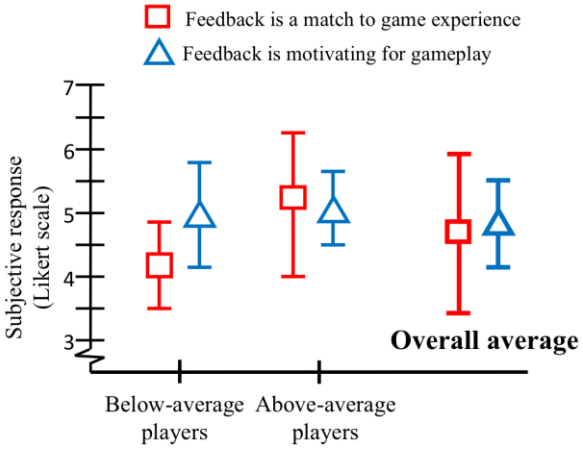


Fig. 4. Likert scale ratings for feedback from players with different skill level. A Likert scale of 1 to 7 (higher is better) is used. Error bars show standard deviation.

The Levene's Test for Equality of Variances, shows that the variability of the two groups is the same. The ANOVA test shows $F(30) = 39.594$, $p = .001$. Because of this, we can conclude that there is statistically significant difference between above-average players and below-average players. Using a Mann-Whitney U-test to examine the participants' subjective rating of how the empathic feedback matched their game experience, we found above-average players ($M = 5.286$, $SD = 1.271$) being significantly different with $p = .027$, $U = 60.0$, $r = -.703$ to the below-average players ($M = 4.273$, $SD = .679$). The above-average players thought the feedback is a better match to their gaming experience than below-average players.

We then used Wilcoxon signed rank test using median equals 4.0 to determine whether there is significant matching of the empathic feedback to their game experience. The above-average players with $p = .001$ has shown a significant effect ($p < .05$) on matching the empathic feedback to their game experience. In contrast, the below-average players with $p = .76$ did not show such significance. However, when we consider all the participants ($M = 4.9375$, $SD = 1.479$), they show a significant effect on matching the feedback to their game experience ($p = .004$). To determine the effect of feedback on motivating the gameplay, we used a Mann-Whitney U-test to examine how empathic feedback improves their gaming experience. We found that the above-average players are not significantly different ($p = .669$, $U = 105.0$) to the below-average players. This indicates that both above-average players and below-average players thought the feedback improves their gaming experience. Using a Wilcoxon signed rank test with median equals 4.0, we show that both above-average players ($M = 5.143$, $SD = 0.526$) with $p = .007$ and below-average players ($M = 5.003$, $SD = 0.848$) with $p = .054$ agree that the empathic feedback significantly improves their gaming experience (figure 4). When we consider all participants ($M = 5.093$, $SD = 0.745$), the overall results indicated that the empathic feedback ($p < .05$) significantly improved their game experience. In summary, the data shows that although all players find the empathic feedback motivating for future matches,

below-average players did not find that the feedback matched their game experience. In contrast, there exists a high match between the feedback and game experience for above-average players.

6 Discussion and Conclusion

In our study, we observe that the three personality types of players are coherent in term of how the empathic feedback affects their gaming experience. This is validated by the study, which shows that the empathic feedback has a significant effect on its match to game experience and also provide motivation to continue gameplay for all extrovert, ambivert and introvert players.

From the analysis on their game performance, participants who are below-average players experience “a narrow view” during the game. Their objective is on winning the game and their focus is on the game dynamics and coordinating their gross motor skills. Thus the feedback does not seem relevant to their gaming experience. On the other hand, participants who are above-average players enjoy the flow of the game experience [5] (fun experience) and become aware of more details of the game, such as graphics and music. They also find the feedback more relevant. In one instance, Bro-cam showed the following feedback with Incongruity humor type to a participant who had skillfully won the game and exhibited ambivert behavior. “You surely suck as a lover, because you’re very good in this game.” The participant responded with a surprise grin and exclaimed “That’s so true, how did you [Bro-cam] know about it!”. Using the body posture of the players as an indicator to their openness and win-lose outcome, we can map different types of humor for participants who exhibited different personality type (i.e. extrovert, ambivert, and introvert) to increase the overall gaming experience. The initial test results show that Bro-cam’s empathic feedback is a promising way to improve the gaming experience.

In conclusion, we contribute a new fully-implemented system that analyzes player posture during gameplay and provides empathic feedback to the player based on the posture "openness" and the win-lose status. With that we try to persuade and motivate players to continue with the game. In addition we provide evidence that this type of system can be "well-received" and "enriches the overall game experience." With our approach we also show that it is possible to easily augment the gaming experience for commercial systems by using cheap commodity hardware (webcam and one additional Kinect). Finally, this setup can be used in subsequent research where it is desirable to study users playing commercial titles, rather than using "toy" games that have been customized for a study.

References

1. Aviezer, H., Trope, Y., Todorov, A.: Body cues, not facial expressions, discriminate between intense positive and negative emotions. Science 338(6111), 1225–1229 (2012)
2. Berkovsky, S., Bhandari, D., Kimani, S., Colineau, N., Paris, C.: Designing Games to Motivate Physical Activity. In: PERSUASIVE 2009 - Persuasive Technology, Fourth International Conference, pp. 37–40. Springer, Heidelberg (2009)

3. Bickmore, T., Gruber, A., Picard, R.: Establishing the computer-patient working alliance in automated health behavior change interventions. Patient Education and Counseling 59, 21–30 (2005)
4. Blumenfeld, R.: Tools and Techniques for Character Interpretation: A Handbook of Psychology for Actors, Writers, and Directors, pp. 115–128. Limelight Editions (2006)
5. Csikszentmihalyi, M.: Flow: The Psychology of Optimal Experience. Harper Perennial, New York (1990)
6. Dekker, A., Champion, E.: Please Biofeed the Zombies: Enhancing the Gameplay and Display of a Horror Game Using Biofeedback. In: Proc. of DiGRA (2007)
7. Eysenck, S., Eysenck, H.J., Barrett, P.: A revised version of the psychoticism scale. Personality and Individual Differences 6, 21–29 (1985)
8. Gamberini, L., Spagnolli, A., Corradi, N., Jacucci, G., Tusa, G., Mikkola, T., Zamboni, L., Hoggan, E.: Tailoring Feedback to Users' Actions in a Persuasive Game for Household Electricity Conservation. In: Bang, M., Ragnemalm, E.L. (eds.) PERSUASIVE 2012. LNCS, vol. 7284, pp. 100–111. Springer, Heidelberg (2012)
9. Gaver, W.: Designing for emotion (among other things). Philosophical Transactions of the Royal Society 364, 3597–3604 (2009)
10. Gilleade, K., Dix, A., Allanson, J.: Affective Videogames and Modes of Affective Gaming: Assist Me, Challenge Me, Emote Me. In: Proc. of DiGRA, pp. 16–20 (2005)
11. Höök, K.: Affective Loop Experiences – What Are They? In: Oinas-Kukkonen, H., Hasle, P., Harjumaa, M., Segerståhl, K., Øhrstrøm, P. (eds.) PERSUASIVE 2008. LNCS, vol. 5033, pp. 1–12. Springer, Heidelberg (2008)
12. Kuikkaniemi, K., Laitinen, T., Turpeinen, M., Saari, T., Kosunen, I., Ravaja, N.: The influence of implicit and explicit biofeedback in first-person shooter games. In: Proc. of CHI 2010, pp. 859–868. ACM Press (2010)
13. Pantic, M., Caridakis, G., André, E., Kim, J., Karpouzis, K., Kollias, S.: Multimodal emotion recognition from low-level cues. In: Cowie, R., Petta, P., Pelachaud, C. (eds.) Emotion-Oriented System, pp. 115–132. Springer, Heidelberg (2011)
14. Nacke, L.E., Kalyn, M., Lough, C., Mandryk, R.L.: Biofeedback game design: using direct and indirect physiological control to enhance game interaction. In: Proc. of CHI 2011, pp. 103–112. ACM Press (2011)
15. Pfeifer, B.: Narrative Combat: Using AI to enhance tension in an Action Game. In: Kirmse, A. (ed.) Game Programming Gems 4, pp. 315–324. Charles River Media (2004)
16. Picard, R.W.: Affective Computing. MIT Press, Cambridge (2000)
17. Picard, R.W.: Affective Computing: Challenges. International Journal of Human-Computer Studies 59(1-2), 55–64 (2003)
18. Strapparava, C., Stock, O., Mihalcea, R.: Computational humour. In: Cowie, R., Petta, P., Pelachaud, C. (eds.) Emotion-Oriented Systems, pp. 609–634. Springer, Heidelberg (2011)
19. Tan, S., Schöning, J., Luyten, K., Coninx, K.: Informing Intelligent User Interfaces by Inferring Affective States from Body Postures in Ubiquitous Computing Environments. In: Proc. of IUI 2013. ACM Press (2013)
20. Tapus, A., Tapus, C., Matarić, M.J.: User-Robot Personality Matching and Robot Behavior Adaptation for Post-Stroke Rehabilitation Therapy. Intelligent Service Robotics Journal, 169–183 (2008)

Trusting Digital Chameleons: The Effect of Mimicry by a Virtual Social Agent on User Trust

Frank M.F. Verberne[1], Jaap Ham[1], Aditya Ponnada[2], and Cees J.H. Midden[1]

[1] Eindhoven University of Technology, Eindhoven, The Netherlands
[2] Indian Institute of Technology,Guwahati, Assam, India

Abstract. Earlier research suggested that mimicry increases liking and trust in other people. Because people respond socially to technology and mimicry leads to increased liking of virtual agents, we expected that a mimicking virtual agent would be liked and trusted more than a non-mimicking one. We investigated this expectation in an automotive setting. We performed an experiment in which participants played an investment game and a route planner game, to measure their behavioral trust in two virtual agents. These agents either mimicked participant's head movements or not. Liking and trust of these virtual agents were measured with questionnaires. Results suggested that for the investment game, mimicry did not increase liking or trust. For the route planner game however, a mimicking virtual agent was liked and trusted more than a non-mimicking virtual agent. These results suggest that mimicry could be a useful tool to persuade users to trust a virtual agent.

Keywords: liking, trust, virtual agent, investment game, route planner game.

1 Introduction

In the near future, we might not need to drive our cars ourselves anymore. In 2010, Google announced that they had developed self-driving autonomous cars, which would not need a human driver anymore [1]. These experimental 'Google cars' have been autonomously driving amongst other human controlled cars on real roads for 140,000 miles with only occasional human intervention and 1,000 miles without any human intervention. Only one accident occurred with one of the cars being rear-ended by another car. With this technology, it is only a matter of time until self-driving autonomous cars will hit the streets of the future.

Although self-driving autonomous cars could have advantages such as safer driving, less congestion, and better fuel-efficiency, the question remains if human drivers will trust technology enough to take over the task of driving their own cars. In the interfaces of such cars, virtual social agents[1] could try to persuade the driver to hand over control. In the current research, we investigate whether mimicry is a useful persuasive technique that can be used by such agents to increase trust in automation technology in cars. For the sake of readability, the term agent(s) will be used in the

[1] In the current paper, we define a virtual social agent as a digital representation of a human, that is controlled by a computer, and not by another human being (the latter being an avatar).

S. Berkovsky and J. Freyne (Eds.): PERSUASIVE 2013, LNCS 7822, pp. 234–245, 2013.

rest of the paper to refer to virtual social agent(s). To be able to persuade drivers to hand over control, it is crucial that they trust the technology.

1.1 Definition of Trust

Trust is necessary in a situation that is characterized by uncertainty and vulnerability [2], which is the case with handing over control to automation technology. In several scientific domains, the concept of trust is measured and used. Although no universally accepted definition of trust exists, a broadly accepted definition of trust has been proposed by Mayer, Davis, and Schoorman [3]:

> Trust is … the willingness of a party to be vulnerable to the actions of another party based on the expectation that the other will perform a particular action important to the trustor, irrespective of the ability to monitor or control that other party. (p. 712)

To gain insight in the effect of mimicry on trust in an agent, we first look at mimicry in human-human interactions.

1.2 The Chameleon Effect

In human-human interactions, humans mimic each other unconsciously on a variety of behaviors including tone of voice [4], facial expressions [5], mood [4], and physical mannerisms [6]. Mimicry has been shown to enhance liking and strengthen bonds between people, even between strangers [6]. This effect has been coined 'the chameleon effect', and is seen as the social glue that binds people together and creates harmonious relationships [7]. Self-other overlap has been suggested to be the mechanism behind mimicry and increased liking [8]. Mimicry has also been shown to increase trust and cooperation behavior in a deal making situation [9]. So mimicry can be used by humans to persuade another human to like and trust him/her more. Could this persuasive technique also be used by agents to achieve the same results?

Research on the media equation hypothesis [10] suggested that people might trust agents in the same way as they trust other humans. Several experiments suggested that humans respond socially to computers, comparable to how humans respond socially to other humans [10]. For instance, people simply like a person more when that person is from the same group (minimal group paradigm; [11]; see also, e.g. [12]). In an experiment where participants interacted with a computer that was presented as a team member or a non-team member, the team member computer was seen as more similar to them and was rated as friendlier than the non-team member computer [13]. Thus, participants socially respond to an artificial non-human team member, comparable to how they respond to a human team member. Collectively, these studies suggest that people apply social rules to computers automatically, even though they are aware that computers are different from humans [10]. Thus, in the definition of trust provided above, the other party does not need to be a human, but can also be an agent.

1.3 The Digital Chameleon

In line with the media equation hypothesis, research suggested that the chameleon effect also takes place when humans are mimicked by agents [14]. In that study, an

agent delivered a persuasive message while either mimicking the head movements of the current participant, or using the recorded head movements of the previous participant (thus not mimicking the current participant). Both liking and persuasiveness of the agent were measured. When the head movements of participants were mimicked, the agent was liked better and was more persuasive than when the head movements were not mimicked. These results support the digital chameleon effect: the chameleon effect for agents.

1.4 The Current Research

As argued, trust is crucial for the acceptance of automation technology. In the current research, we argue that mimicry can be used as a persuasive technique by agents to be more trusted by their user. Just as mimicry influences the trust people have in other humans, we expect that mimicry also influences the trust people have in agents.

More specifically, our main hypothesis is that a mimicking agent will be trusted (on a behavioral and cognitive level) and liked more than a non-mimicking one. Furthermore, we expect that liking mediates this increase in trust. That is, we expect that a mimicking agent will be liked more and *therefore* be trusted more.

Our exploratory hypothesis is that the self-other overlap resulting from mimicry is a moderator for the digital chameleon effect. That is, people tend to like a mimicking agent more, because it is more similar to them. This effect should be most profound for people with positive explicit/implicit self-esteem. Therefore, we predict that a mimicking agent will be liked and trusted more, and that the more self-other overlap, the more pronounced the effects of mimicry will be.

2 Methods

2.1 Participants and Design

Forty participants (17 women and 23 men, average age = 21.9 years, *SD* = 2.4) were randomly assigned to the conditions of a single factor (mimicry: mimicked vs. non-mimicked) between subject design. Dependent measures consisted of two behavioral trust measures, a trust questionnaire and a liking questionnaire. The experiment lasted approximately 30 minutes, for which participants were paid €5 plus their bonus earned by playing the investment game and route planner game. One participant scored as an outlier (based on the criteria of [15]) on the investment game and the liking questionnaire, and was therefore excluded from data analysis, leaving a sample size of 39.

2.2 Materials and Measures

Tracking. An Ascension Flock of Birds™ 6DOF magnetic-field position sensor was attached to a cap for tracking the orientation of participants' heads. While wearing this cap, participants could freely move their head during the experiment. The orientation data was used for mimicking (see below) and was recorded to be used in the non-mimicked condition.

Agents. Two male agents (see Appendix A) were used to play the investment game and route planner game with (see dependent measures). Two agents were used so that the investment game agent was different from the route planner game agent, and experience with the one would not affect reactions to the other. Both agents were from the Vizard Complete Character set. The order of the two agents was counterbalanced between participants. For one half of the participants, the investment game agent was the left agent of Appendix A and the route planner game agent the right one, for the other half, vice versa. In a pretest, we measured the likeability and trustworthiness of both agents. Results suggested that the agents did not differ in how likeable and trustworthy they were perceived. Both agents spoke with the same synthetic Dutch voice and blinked their eyes at a natural rate. While speaking, they moved their lips in synch with the speech. Other than the head movements, eye blinking, and lip movements, the agents did not move. The investment game agent was called Jeroen and the route planner game agent was called Max. These names will be used in the rest of the paper to refer to these agents.

Agent's Head Movements. In the mimicked condition, both Jeroen and Max mimicked the head movements of the participant (yaw, pitch, and roll; which are the rotations about the Y-, X-, and Z-axis respectively), with a delay of 4 seconds (in line with [14]). Head movements were mirror mimicked, so that when a participant turned his head to the left, Jeroen and Max would turn their head to their right (like your own mirror image would do). In the non-mimicked condition, they moved their head using the recorded head movements of the previous participant (in line with [14]).

Dependent Measures. To test our main hypothesis, we included the dependent measures described below.

Investment Game Decision. Behavioral trust was measured with the investment game [16]. Participants and agent Jeroen were both given 10 credits. Participants had to choose how many credits to give to Jeroen. Every credit given was tripled, and Jeroen would then decide how many credits to give back to the participant. To keep the situation uncertain, his decision was revealed to the participants at the very end of the experiment. The number of credits participants decided to give to Jeroen is the behavioral measure of trust in this game. The more credits are given, the more the participant trusted him.

Route Planner Game Decision. Behavioral trust was also measured using a route planner game, in which the rules were similar to the investment game. At the beginning of the route planner game, participants were given 20 credits. Agent Max represented a navigation system, and participants were given 10 routes to plan. For each route, participants had two choices: either plan the routes themselves, or let Max plan the route for them. If participants chose to plan the route themselves, they lost one credit. If participants chose to let Max plan the route, they lost two credits. To keep the situation uncertain, the routes planned by Max were shown at the end of the experiment. If the planned route happened to be the fastest route (ostensibly according to a database), participants would get four credits back for that route. Every route given to Max could lead to a potential gain of three credits (like in the investment game). The amount of routes participants decided to give to Max is the behavioral measure of trust in this game. The more routes are given, the more participants trusted him.

The investment game and the route planner game measure a different kind of behavioral trust. In the investment game, *relational trust* is measured: participants can earn credits based on the *intention* of the agent. In the route planner game, it is about *calculative trust*: participants can earn credits based on the *competence* of the agent. Both the investment game and the route planner game involved real monetary stakes. Every credit participants had left at the end of the experiment was worth €0.05. Although the exact monetary value of a credit was unknown to the participants, they were told that the credits left at the end would determine their monetary bonus.

Trust. Trust in Jeroen and Max (on a cognitive level) was measured by a questionnaire [17] with twelve questions with a seven-point Likert scale (1 = "totally disagree", 7 = "totally agree"). Answers to these questions were averaged to form a reliable measure of trust (Cronbach's Alpha's = .89 & .87, for Jeroen and Max respectively). Higher scores indicate more trust.

Liking. As in [14], liking of Jeroen and Max was measured on 13 different dimensions using the partner ratings items of [18]. These dimensions include approachable, confident, likable, interesting, friendly, sincere, warm, competent, informed, credible, modest, honest, and trustworthy. Liking of the agent was measured on a seven-point Likert scale (1 = "totally disagree", 7 = "totally agree"). Answers were averaged to form a reliable measure of liking (Cronbach's Alpha's: .87 & .83, for Jeroen and Max respectively). Higher scores indicate more liking.

Exploratory Measures. We included several exploratory measures to explore the effect of mimicry on self-other overlap and the role of self-esteem. Furthermore, because trust is necessary in a situation characterized by uncertainty and vulnerability [2], we also included measures to explore whether mimicry affects experienced uncertainty and vulnerability during the investment game and route planner game.

Self-other Overlap. Self-other overlap was measured using the Inclusion of the Other in the Self scale [19]. We used two versions: a pictorial and a continuous one. Both versions were used to measure self-other overlap with both agents. In both versions, two circles were presented, one representing the participant and one representing the agent. In the *pictorial* version, seven pairs of overlapping circles were presented (ranging from no overlap to a lot of overlap) and participants had to choose which pair depicted their relationship with the agent best. In the *continuous* version, participants could click the circle representing themselves and drag it closer to the circle representing the agent, creating as much (or little) overlap as desired to indicate their relationship with the agent.

Explicit Self-esteem. Explicit self-esteem was measured using the Rosenberg Self-Esteem Scale [20]. This scale has 10 four-point Likert scale items, ranging from 'strongly agree' to 'strongly disagree'. Answers to these items were averaged to form a reliable measure of explicit self-esteem (Cronbach's Alpha: .87). Higher scores indicate more explicit self-esteem.

Implicit Self-esteem. Implicit self-esteem was measured using a self-esteem Single-Target Implicit Association Test (STIAT; [21]). In this STIAT, participants had to categorize self-relevant (the words me, my, myself, and their own first and last name), positive, and negative words as either positive or negative. In one block, self-relevant

and *positive* targets shared the same response key; in another block, self-relevant and *negative* targets shared the same response key. Positive implicit self-esteem is indicated when reaction times in the self+positive block are shorter than in the self+negative block. The bigger the difference, the higher the implicit self-esteem.

Expected Credits. For both the investment game and the route planner game, a single item was used to measure expected gains with the question "How many credits do you expect to get from the game you've just played?".

Perceived Risk. Perceived risk of Jeroen and Max was measured by a questionnaire with five questions (e.g. "I think the choice to give (Jeroen credits/Max routes) was risky.") with a seven-point Likert scale (1 = "totally disagree", 7 = "totally agree"). Answers to these questions were averaged to form a measure of perceived risk. However, these questions appeared only to be reliable for Max (Cronbach's Alpha's = .57 & .73, for Jeroen and Max respectively). Responses were coded such that higher scores indicate more risk.

Perceived Competence. Perceived competence of Jeroen was measured with three questions (e.g. "I think Jeroen is intelligent.") with a seven-point Likert scale (1 = "totally disagree", 7 = "totally agree"). Perceived competence of Max was measured with seven questions (e.g. "Max can plan routes well."). Answers to these questions were averaged to form a reliable measure of perceived competence. (Cronbach's Alpha's = .89 & .67, for Jeroen and Max respectively). Responses were coded such that higher scores indicate more competence.

IQ. For both agents, we measured participants' estimation of their IQ. This was a single item with the question "The average IQ of a human is 100. How high would you estimate the IQ of Jeroen/Max?".

2.3 Procedure

Participants were recruited using the JF Schouten participant database of Eindhoven University of Technology. When participants entered the lab individually, they were seated in front of a television screen (at approximately 2 meters distance from a 47" Samsung Full HD TV) and were instructed to put on a cap equipped with the tracking sensor. Next, Jeroen appeared on the screen and verbally explained the rules of the investment game (lasting 102 seconds). During this explanation, Jeroen either mimicked the participant or not, depending on condition. After finishing the investment game, participants had to complete the trust, liking, expected credits, self-other overlap, perceived risk, perceived competence, and IQ measures for Jeroen. When they completed those measures, participants were taken to another room for a few minutes to complete a non-related distraction task. Then, participants were seated again and were put on the cap. Max appeared on the screen and verbally instructed them how to play the route planner game. During this instruction and while playing the game, Max either mimicked the participants or not, depending on condition. After finishing the route planner game, participants had to complete the trust, liking, expected credits, self-other overlap, perceived risk, perceived competence, and IQ measures for Max. Next, explicit and implicit self-esteem were measured. Then, if any amount of routes were given to Max to plan, results of those routes were shown.

Then, participants were shown the amount of credits they had left after playing the investment game and route planner game, which determined their bonus pay for the experiment. Lastly, participants were debriefed, paid and thanked for their participation.

3 Results

3.1 Investment Game

Effect of Agent. To rule out that participants responded differently towards the two agents used in the investment game, a one-way MANOVA was conducted with agent as the independent variable and investment game decision, trust, and liking as dependent variables. No significant effects were revealed (all p's > .250). Therefore, we conclude that participants responded similar to both agents in the investment game.

Main Analysis. To test our main hypothesis, a one-way MANOVA was conducted with mimicry as the independent variable and investment game decision, trust, and liking as dependent variables. Results revealed no significant multivariate effect of mimicry, $F(3,35) = 2.65$, $p = .500$, $\eta_p^2 = .06$. All main effects of mimicry were non-significant (all p's > .131). This analysis did not provide evidence for our main hypothesis. That is, results did not suggest that a mimicking agent was liked and trusted more than a non-mimicking one.

Exploratory Analysis. Separate one-way ANOVA's were conducted with mimicry as the independent variable and expected credits, self-other overlap (pictorial and continuous), perceived risk, perceived competence, and IQ as dependent measures. A main effect of mimicry on IQ was revealed, $F(1, 38) = 5.89$, $p = .020$, $\eta_p^2 = .14$. Participants rated the IQ of the agent in the mimicked condition to be lower (M = 98.00, SE = 3.65) than in the non-mimicked condition (M =110.68, SE = 3.74). All other effects were non-significant (all p's > .208). This analysis did not provide evidence for our exploratory hypothesis. That is, results did not suggest that self-other overlap is a mediator for the digital chameleon effect: mimicry did not increase self-other overlap.

3.2 Route Planner Game

Effect of Agent. To rule out that participants responded differently towards the two agents used in the route planner game, a one-way MANOVA was conducted with agent as independent variable and route planner decision, trust, and liking as dependent variables. No significant effects were revealed (all p's > .301). Therefore, we conclude that participants responded similar to both agents in the route planner game.

Main Analysis. To test our main hypothesis, a one-way MANOVA was conducted with mimicry as the independent variable and route planner decision, trust, and liking as dependent variables. In line with our hypothesis, results revealed a marginally significant multivariate effect of mimicry, $F(3,35) = 2.65$, $p = .064$, $\eta_p^2 = .19$. The main

effect of mimicry on route planner decision was significant, $F(1,38) = 4.83$, $p = .034$, $\eta_p^2 = .12$. Participants gave more routes to the agent in the mimicked condition ($M = 5.50$, $SE = 0.53$) than in the non-mimicked condition ($M = 5.05$, $SE = 0.55$). The main effect of mimicry on trust was marginally significant, $F(1,38) = 2.61$, $p = .058$ (1-tailed), $\eta_p^2 = .07$. In the mimicked condition, the agent was trusted more ($M = 4.40$, $SE = 0.20$) than in the non-mimicked condition ($M = 4.08$, $SE = 0.20$). The main effect of mimicry on liking was significant, $F(1,38) = 5.29$, $p = .027$, $\eta_p^2 = .13$. Participants liked the agent more in the mimicked condition ($M = 4.63$, $SE = 0.15$) than in the non-mimicked condition ($M = 4.29$, $SE = 0.16$). This analysis did provide evidence for our main hypothesis. That is, results did suggest that a mimicking agent was liked and trusted more than a non-mimicking one.

Mediation Analysis. To test whether increased liking mediated the effect of mimicry on the route planner decision, a mediation analysis (following the steps of [22]) was conducted to reveal the direct (Path c) and the indirect effects (Paths a and b) of mimicry on the route planner decision. A Sobel test [23] showed that the indirect effect was marginally significant (Sobel $z = 1.41$, $p = .08$ (1-tailed)). The initial effect of mimicry on route planner decision (Path c) becomes non-significant after controlling for liking (Path c'; see Appendix B), which shows that liking mediates the initial effect. The results from this analysis confirm our expectations that liking, resulting from mimicry, mediates the effect of mimicry on route planner decision.

Exploratory Analysis. Separate one-way ANOVA's were conducted with mimicry as the independent variable and expected credits, self-other overlap (pictorial and continuous), perceived risk, perceived competence, and IQ as dependent measures. Again, a main effect of mimicry on IQ was revealed, $F(1, 38) = 4.27$, $p = .046$, $\eta_p^2 = .10$. Participants rated the agent in the mimicked condition less intelligent ($M = 97.90$, $SE = 3.45$) than in the non-mimicked condition ($M = 108.11$, $SE = 3.54$). Furthermore, a main effect of mimicry on expected credits was revealed, $F(1, 38) = 6.59$, $p = .014$, $\eta_p^2 = .15$. In the mimicked condition, participants expected to get more credits back from the route planner game ($M = 14.45$, $SE = 0.54$) than in the non-mimicked condition ($M = 12.47$, $SE = 0.55$). All other effects were non-significant (all p's $> .136$). This analysis did not provide evidence for our exploratory hypothesis. That is, results did not suggest that self-other overlap is a mediator for the digital chameleon effect: mimicry did not increase self-other overlap.

4 Discussion and Conclusion

The current research investigated whether an agent could persuade participants to trust and like him more by mimicking them. To investigate this question, participants played an investment game and a route planner game with two agents that either mimicked their head movements, or did not. Non-mimicking agents moved their head using the recorded head movements of the previous participant. Results only partly provided support for our hypothesis. That is, for the investment game, results did not suggest that an agent could persuade participants to like and trust it more by mimicking them. For the route planner game however, results suggested that participants liked and trusted a mimicking agent more than a non-mimicking one.

Furthermore, as expected, the effect of mimicry on trust behavior was mediated by liking.

Overall, results suggest some support for our main hypothesis. In the route planner game, the predicted effects were found: participants liked and trusted (on a behavioral and cognitive level) a mimicking agent more than a non-mimicking one. However, in the investment game, we did not find support for our main hypothesis. We argue that there are two plausible explanations as to why the results from the investment game did not support our main hypothesis.

The first, methodological, explanation is that the order of both games was always the same: the investment game was played first. Mimicry could take some time to lead to positive effects, and before deciding in the investment game, participants were exposed to 102 seconds of mimicry in our experiment (versus 195 seconds in [14]). Our duration of mimicry could have been below a certain threshold for positive effects to emerge in the investment game. Future research could identify this specific threshold of mimicry (if there is any), in either a human-human interaction or a human-agent interaction. To be on the safe side, the duration of mimicry should be at least 195 seconds in future research before measuring trust behavior.

The second, theoretical, explanation is that the investment game and the route planner game measure a different kind of behavioral trust. As argued before, in the investment game, *relational trust* is measured: participants can earn credits based on the *intention* of the agent. In the route planner game, it is about *calculative trust*: participants can earn credits based on the *competence* of the agent (see [24] for a detailed discussion of the distinction between relational and calculative trust). In other words, the investment game choice was risky because participants did not know how many credits Jeroen was willing to give back. The route planner game choice was risky because participants did not know how well Max could plan routes for them. Our results suggest that mimicry seems to increase calculative trust in an agent, but not relational trust. Possibly, people do not easily attribute willingness to an agent, explaining why mimicry does not increase relational trust in an agent. Future research could further disentangle the effect of mimicry on both calculative and relational trust by using different measures. In this study, we tried to do just this by using both the investment game and the route planner game. Future research could explore different paradigms with similar differences.

Results do not support our exploratory hypothesis. That is, results do not suggest that the self-other overlap resulting from mimicry is a moderator for the digital chameleon effect. Although mimicry has been shown to lead to increased *neural* self-other overlap [8], we failed to detect an increase of self-other overlap due to mimicry using pictorial measures in our study. Future research could focus on different ways of measuring self-other overlap more directly to further investigate this exploratory hypothesis. In addition, the role of explicit/implicit self-esteem in combination with self-other overlap could be tested in future research.

Unexpectedly, participants rated both agents less intelligent when they mimicked them versus when the agents used recorded head movements. One possibility is that mimicry decreases IQ. Research has shown that positive effects not only depend on the ability to mimic, but also on the ability *not* to mimic [25]. Using our mimicry algorithm, participants were mimicked precisely after a delay of 4 seconds, and every rotation of the head was mimicked. The algorithm therefore resulted in consistent but

rigid mimicry. When humans mimic each other, mimicry is more *flexible*. Humans can choose what movements to mimic, and also when *not* to mimic. Future research could focus on improving mimicry algorithms, to make them model natural mimicry between humans more closely than current algorithms.

Another possibility is that the wording of our question explains the results. We compared the agent to *humans* in our question. Because the voice of the agent was synthetic, and the appearance of the agent was clearly computer generated, comparing the IQ of the agent to the IQ of humans in general might lead to downward comparisons. Mimicry could have made the agent more humanlike, but its non-human attributes (voice and appearance) could have decreased its estimated IQ. Although unexpected, future research could further explore the effect of mimicry on IQ estimations of agents.

In the current study, a specific form of mimicry was used, namely mimicking head movements. As agents become more advanced and more capable of mimicking a wider variety of human behavior, future research could test the effect of other types of mimicry by agents. These could include mimicking the voice [4], the facial expressions [5], and even the mood of the user [6]. We would expect similar effects of these types of mimicry by agents. That is, we do not suggest that mimicking head movements uniquely increases liking and trust of agents. Furthermore, we do not expect that the positive effects of mimicry only occur in an automotive context. Mimicry has been shown to have positive effects in a wide variety of human-human interaction settings.

In conclusion, the current study suggested some evidence that mimicry can be used by agents to persuade their user to like and trust them more. That is, results suggested that mimicry increased liking and trust behavior towards an agent, when the trust situation was competence based. This effect should be replicated in future studies to improve the strength of the evidence. Nevertheless, these results take a first step in making agents inside a self-driving car more trustworthy to be able to persuade drivers to give up control of driving to their cars.

References

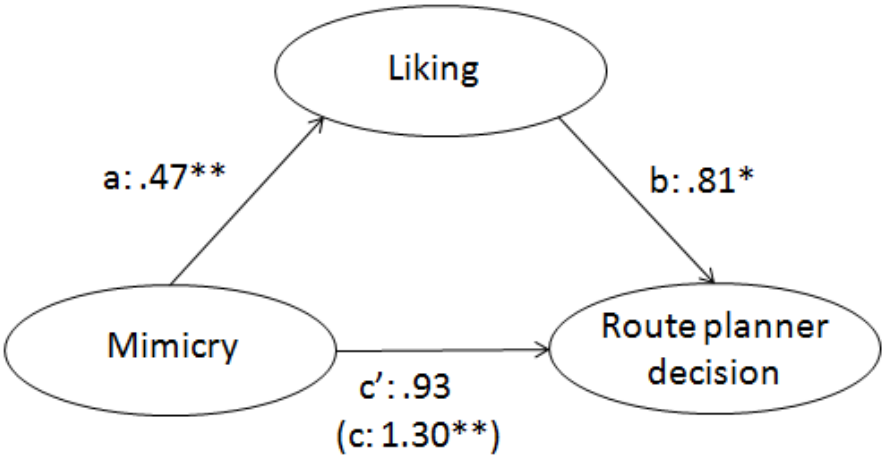
1. Markhof, J.: Google cars drive themselves, in traffic. The New York Times (October 9, 2010), `http://www.nytimes.com` (retrieved)
2. Lee, J.D., See, K.A.: Trust in automation: Designing for appropriate reliance. Human Factors 46, 50–80 (2004)
3. Mayer, R.C., Davis, J.H., Schoorman, F.D.: An integrative model of organizational trust. Academy of Management Review 20, 709–734 (1995)
4. Neumann, R., Strack, F.: Mood contagion: The automatic transfer of mood between persons. Journal of Personality and Social Psychology 78, 211–223 (2000)
5. Hsee, C.K., Hatfield, E., Carlson, J.G., Chemtob, C.: The effect of power on susceptibility of emotional contagion. Cognition and Emotion 4, 327–340 (1990)
6. Chartrand, T.L., Bargh, J.A.: The Chameleon effect: The perception-behavior link and social interaction. Journal of Personality and Social Psychology 76, 893–910 (1999)
7. Lakin, J.L., Jefferis, V.E., Cheng, C.M., Chartrand, T.L.: The chameleon effect as social glue: Evidence for the evolutionary significance of nonconscious mimicry. Journal of Nonverbal Behavior 27, 145–162 (2003)

8. Inzlicht, M., Gutsell, J.N., Legault, L.: Mimicry reduces racial prejudice. Journal of Experimental Social Psychology 48, 361–365 (2012)
9. Maddux, W.W., Mullen, E., Galinsky, A.D.: Chameleons bake bigger pies and take bigger pieces: Strategic behavioral mimicry facilitates negotiation outcomes. Journal of Experimental Social Psychology 40, 461–468 (2008)
10. Reeves, B., Nass, C.: The media equation: How people treat computers, television, and new media like real people and places. Cambridge University Press, New York (1996)
11. Tajfel, H.: Experiments in intergroup discrimination. Scientific American 223, 96–102 (1970)
12. Turner, J.C., Brown, R.J., Tajfel, H.: Social comparison and group interest in ingroup favoritism. European Journal of Social Psychology 9, 187–204 (1979)
13. Nass, C., Fogg, B.J., Moon, Y.: Can computer be teammates? International Journal of Human-Computer Studies 45, 669–678 (1996)
14. Bailenson, J., Yee, N.: Digital chameleons: Automatic assimilation of nonverbal gestures in immersive virtual environments. Psychological Science 16, 814–819 (2005)
15. Tukey, J.W.: Exploratory data analysis. Addison-Wesley, Reading (1977)
16. Berg, J., Dickhaut, J., McCabe, K.: Trust, reciprocity, and social history. Games and Economic Behavior 10, 122–142 (1995)
17. Jian, J., Bisantz, A.M., Drury, C.G.: Foundations for an empirically determined scale of trust in automated systems. International Journal of Cognitive Ergonomics 4, 53–71 (2000)
18. Guadagno, R.E., Cialdini, R.B.: Online persuasion: An examination of gender differences in computer-mediated interpersonal influence. Group Dynamics: Theory, Research, and Practice 6, 38–51 (2002)
19. Aron, A., Aron, E.N., Smollan, D.: Inclusion of other in the self scale and the structure of interpersonal closeness. Journal of Personality and Social Psychology 63, 596–612 (1992)
20. Rosenberg, M.: Society and the adolescent self-image. Princeton University Press, Princeton (1965)
21. Bluemke, M., Friese, M.: Reliability and validity of the Single-Target IAT (STIAT): Assessing automatic affect towards multiple attitude objects. European Journal of Social Psychology 38, 977–997 (2008)
22. Baron, R.M., Kenny, D.A.: The moderator-mediator variable distinction in social psychological research: Conceptual, strategic and statistical considerations. Journal of Personality and Social Psychology 51, 1173–1182 (1986)
23. Sobel, M.E.: Asymptotic confidence intervals for indirect effects in structural equation models. In: Leinhart, S. (ed.) Sociological Methodology 1982, pp. 290–312. Jossey-Bass, San Francisco (1982)
24. Earle, T.C.: Distinguishing trust from confidence: Manageable difficulties, worth the effort. Risk Analysis 30, 1025–1027 (2010)
25. Kavanagh, L.C., Suhler, C.L., Churchland, P.S., Winkielman, P.: When it's an error to mirror: The surprising reputational costs of mimicry. Psychological Science 22, 1274–1276 (2011)

Appendix A

The two male virtual social agents used in the experiment. For half of the participants, the investment game agent was the left one, and the route planner game agent the right one. For the other half, vice versa. In the experiment, both agents moved their head (either mimicking the participant or not), blinked their eyes, and moved their lips in synch with synthetic speech.

Appendix B



Mediation model depicting the coefficients (Bs) of the direct (Path c) and indirect (Paths a and b) effects of condition on route planner decision. Although the direct effect is significant, this effect becomes non-significant after adding liking as a mediator. * $p < .05$ (1-tailed) ** $p < .05$.

Toward the Design of a Dashboard to Promote Environmentally Sustainable Behavior among Office Workers

Ray Yun[1], Bertrand Lasternas[1], Azizan Aziz[1], Vivian Loftness[1], Peter Scupelli[2], Anthony Rowe[3], Ruchie Kothari[1], Flore Marion[1], and Jie Zhao[1]

[1] School of Architecture,
[2] School of Design,
[3] Electrical and Computer Engineering,
Carnegie Mellon University
{ryun,azizan,scupelli,loftness,jayzhao}@cmu.edu,
{blastern,ruchiek,fmarion}@andrew.cmu.edu,
agr@ece.cmu.edu

Abstract. In the United States, over three billion dollars are spent due to office equipment being left on when not in use during the weekend and at night. There is very little incentive for office workers to save energy because utility bills are not directly their responsibility. Our goal is to find ways to reduce the negative impact of this pervasive phenomenon by applying persuasive technologies to create awareness and encourage office workers towards more environmentally sustainable behavior. To this end, we conducted a literature review to investigate the persuasive methods appropriate to the field of building controls. We then proceeded to develop "dashboard-controllers" that enable office workers to control energy-using components with expert feedback to save energy.

1 Introduction

A single US company with 10,000 computers annually spends $260,000 and contributes 1.871 tons of CO_2 to climate change due to the computers that are left on during the night and the weekend [1]. Office workers' habitual behavior can waste a considerable amount of money and seriously damage the environment.

In the field of Persuasive technology, various innovative systems have been invented to motivate people to change their behavior in a sustainable way. For example, ambient display [e.g., 2] can subtly provide feedback regarding a user's energy consumption in real time, computer games [e.g., 3] can educate people on how to save energy in an enjoyable manner, mobile or web applications [e.g., 6] can provide virtual rewards based on a user's sustainability performance. However, as Foster et al. pointed out [7], many of the studies targeted residential areas, but few studies focus on commercial buildings. Behavioral research on sustainability in commercial buildings is critical. Commercial buildings account for 36% of all electricity usage and 18% of carbon emissions [8], a significant opportunity to conserve energy. In addition,

S. Berkovsky and J. Freyne (Eds.): PERSUASIVE 2013, LNCS 7822, pp. 246–252, 2013.

office workers are oftentimes unaware of the environmental impact of their behavior and care little about saving energy because they are not responsible for energy-related bills. In this paper, we introduce our literature review, preliminary studies and then present our dashboard prototype to promote environmentally sustainable practices in the workplace.

2 Intervention Techniques for Sustainability

We reviewed literature in the field of Persuasive technology, HCI, and environmental psychology that touch on three topics: the environment, behavior change, and the workplace [11]. We summarized the strategies, aggregated the examples and design issues, and categorized them into nine intervention techniques.

Education: This teaches why people should conserve energy and how they can do it. Education can contribute to 28% of energy savings [16]. Today, leading companies put effort into building websites for online training, campaigns, and sharing environmental tips [18].

Advice: This gives suggestions about what people can do to reduce energy consumption. The form of advice varies but may include emails, text messages, and an agent's dialogue. Advice can contribute to a 14.4% reduction of electricity consumption [13] and personalized advice can increase pro-environmental behavior [19].

Self-monitoring: This allows people to learn their energy consumption patterns, and explore where they could save energy. Self-monitoring alone can reduce electricity consumption up to 7% [10] and the rate can be improved when the monitoring shows appliance-specific data [13] frequently and immediately [15], compared with historic data [13] in monetary units [7].

Comparison: This allows people to compare their performance to others. Foster [7] showed that office workers could be effectively motivated if they realize what others do for sustainability. The study of Siero et al. [12] shows employees achieve more energy savings when they are provided with comparative information on energy consumption. A case study shows that one university could save $14,000 in seven weeks by using a website that provides comparative data [9].

Control: This provides easy and simple ways to control energy consumption to save energy. This is similar to Fogg's reduction principle [15], which says "reducing complex behavior into simple tasks increases the cost/benefit ratio of the behavior and motivates users to perform it." With this approach, up to 55% of energy consumption can be saved [14], but it still has not been thoroughly studied.

Reward: A reward is a prize for achieving target behavior. Monetary rewards can make up to a 15% reduction of electricity usage and larger rewards promote pro-environmental behavior better [17]. Foster et al. [7] shows that office workers would prefer realistic and tangible rewards rather than virtual ones.

The first four techniques are informational interventions, which are effective to change one's attitude and intention, and the last two are structural interventions, which change a behavior cost and circumstances [4]. If the anti-environmental behavior has already become one's habit, only informational interventions alone may not be enough [5], and structural interventions can be more effective to influence people's behavior. There are three more principles reviewed here, goal setting, engagement and communication but they are currently out of the research scope and will be included in the future studies.

3 Pilot Studies

We conducted several feature prioritization studies to see what types of information or features can affect people's attitude and lead to energy conservation. The results from these studies were incorporated in a web-based dashboard-controller that monitors the energy consumption of office appliances. The dashboard also enables office occupant control of each appliance. We recruited 22 people at three sites (6 people in a university lab, 8 people in a university office, and 8 people in a government research lab. No control group was formed in this pilot study due to the sample size of the groups.) More than 120 appliances were monitored for 3 to 8 weeks. Based on the collected data, we created energy profiles and patterns for each appliance and identified three thresholds of "Active," "Idle," and "Standby". After baseline data collection, the dashboard interface was installed for each user. A user's interactions (e.g., which buttons they clicked) are logged to understand what features are used.

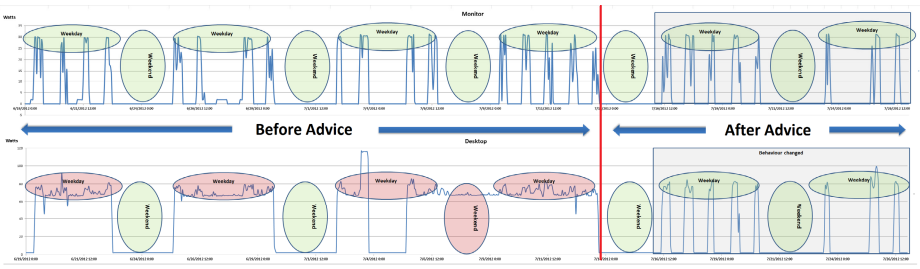


Fig. 1. Awareness impact on User's behavior

3.1 Education (Awareness) and Advice

Fig. 1 represents one of our users' consumption data of her computer screen and desktop computer over six weeks. The chart at the top of the Fig. 1 shows the monitor consumption and the chart at the bottom shows the same person's computer consumption during the same period. According to the chart, the monitor was correctly set up and turned off when not in use but the computer was not (the pink highlighted area in the bottom chart). The user was informed of the mismatch between the monitor and computer, and it was suggested that she change the computer power management setting for energy efficiency. The recommendation came as a surprise because she

thought the computer was in sleep mode when the monitor was off, but it turned out that the computer was still in idle mode. She immediately followed the recommendation and reduced her energy consumption by more than 70% (equivalent to $40 over a year).

3.2 Self Monitoring and Comparison

After the baseline study, we installed our dashboard prototype for one of the groups and allowed the users to monitor their energy consumption and compare it with others. One of the users realized he uses too much energy compared to others and identified his antiquated server computer as the culprit. By replacing the outmoded server with an Energy Star model, the new computer saves more than 500 kWh a year (equivalent to $65), and consumes only 35% of the energy compared to the decommissioned server computer.

4 Dashboard

Based on the intervention techniques we reviewed and the findings from the pilot study, we designed a dashboard targeting office workers. The web-based dashboard incorporates features that are likely to motivate behavior change (Fig. 2).

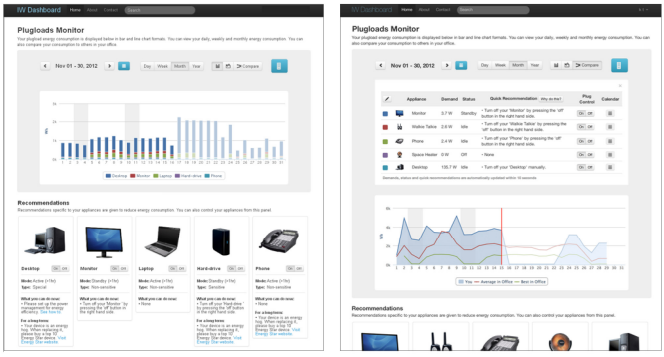


Fig. 2. Self-monitoring and recommendation (left), and control panel and comparison chart (right)

4.1 Design Focus

Self-monitoring and Comparison: Based on the literature review, we designed the chart to display real-time and historic data for individual devices [13,15]. The chart presents data with different intervals (day, week, month, year) and chart types (bar, area, line) to give users various viewpoints of their energy consumption. The chart also displays past data, so users can understand what their consumption was (e.g. a week ago) and predict their future performance. In the comparison section, the dashboard provides the average consumption of similar users and the best practices in the group.

Recommendations: The dashboard provides short-term and long-term recommendations. The short-term recommendations suggest what users can easily do immediately and are generated based on the decision tree that we created using each device's usage status (active/idle/sleep/off), how long the device has been in that status, and the device's sensitivity (sensitive/non-sensitive/special). It suggests to not only turn off the specific item when not in use, but also to set up power management, a one-time action. Long-term recommendations may suggest the replacement of an energy hog for a more energy-efficient device such as an energy-star product to eventually save energy for the long-term.

Control: For office workers to safely control their items with the dashboard, we identified the sensitivity of typical devices in the workplace. For example, we categorized the following items as sensitive: 1) main desktop computers and monitors to access the dashboard, 2) servers, modems, routers to enable network connectivity, and 3) refrigerators that should be turned on all the time. All the rest are categorized as controllable items (e.g., desk lamps, secondary monitor). The dashboard allows users to control the items individually or as a group.

Reward: At the end of every week the dashboard analyzes each user's performance and sends out a weekly report to them. The report includes personalized recommendations based on the user's performance in regards to energy consumption. The users are also reminded that the top 50% of users will be entered in a raffle to win $75 gift cards.

4.2 System Structure

The Plugwise (http://plugwise.com) devices are chosen to monitor appliances' energy consumption. The Plugwise circles are connected to users' appliances and send the electricity consumption data wirelessly to the main server. Every second the server automatically sends the data to update the database in a web server. All the content and visualization in the dashboard is created based on the data in the web server. When users control their items (turn them on or off), the main server detects it and sends a command to the circles to actuate the device. Recommendations and comparative data are also generated in real-time in the web server.

4.3 Design Issues

Here we will highlight several design issues to improve our dashboard. First, control of shared appliances has not been implemented yet, but we see it as another great opportunity to save energy. Based on our six-week baseline study with twenty-two participants, more than 90% of the shared appliances were left on overnight and during the weekend. This accounts for more than 25 % of total energy consumption. The responsibility of all these appliances will be distributed to the office users through the dashboard. We believe that automated control is also a powerful feature. We are now implementing calendar and occupancy-based control features. The calendar feature allows people to schedule when to turn on or off their appliances and occupancy-based control turns the devices on or off based on the user's occupancy in the office. From our baseline study, 40% of the total energy can be saved with this feature.

Lastly, we are currently in the first stage of evaluating our persuasive system and comparing consumption data with the dashboard and without the dashboard. The next stage is to break down the persuasive factors and evaluate them individually and as groups to see which persuasive factors can motivate office workers effectively.

4.4 Preliminary Results

After one month into the study, we conducted initial data analysis and a questionnaire survey in the middle of the study. Two of the sites showed significant energy savings (university lab:31.5%, government research lab:-5%, university office:30%). To understand what contributed to this result, we gave questionnaires about the dashboard's learnability, usability and engagement to our users and eleven of them answered. Most of them think they received useful and clear information (10/11) and it influenced them to behave environmentally (8/11). They enjoyed receiving and reading the weekly reports (8/11). Self-monitoring and comparison interventions were appreciated the most (6/11,5/11, multiple-choice question), and control intervention the least (2/11). The workers in the government lab had an internal policy to keep on the lab computers all the time, so our recommendations did not work well for one of their most power-consuming devices.

5 Conclusion

The results from our pilot study shows that up to 40% overall energy savings can potentially be targeted and one of the sites (the university office) achieved 54% and 79% saving during the weekend and weekday nights respectively through the effectiveness of using persuasive methods to engage occupants in pro-environmental behavior. We will revise the intervention techniques based on this study to move towards a more persistent contribution to the maximum energy savings and conduct longitudinal studies with a larger sample size in the near future.

Acknowledgments. This study is supported by the Greater Philadelphia Innovation Cluster for Energy Efficient Buildings project (GPIC, a U.S. DOE Innovation Hub, Subtask 6.4) under U.S. Department of Energy Award Number EE0004261.

References

1. 1E PC Energy Report: PC Energy Report 2009, USA, United Kingdom, Germany (2009), http://www.1e.com/energycampaign/downloads/PC_EnergyReport2009-US.pdf
2. Kim, T., Hong, H., Magerko, B.: Designing for Persuasion: Toward Ambient Eco-Visualization for Awareness. In: Ploug, T., Hasle, P., Oinas-Kukkonen, H. (eds.) PERSUASIVE 2010. LNCS, vol. 6137, pp. 106–116. Springer, Heidelberg (2010)
3. Bang, M., Gustafsson, A., Katzeff, C.: Promoting New Patterns in Household Energy Consumption with Pervasive Learning Games. In: de Kort, Y.A.W., IJsselsteijn, W.A., Midden, C., Eggen, B., Fogg, B.J. (eds.) PERSUASIVE 2007. LNCS, vol. 4744, pp. 55–63. Springer, Heidelberg (2007)

4. Steg, L., Vlek, C.: Encouraging pro-environmental behaviour: An integrative review and research agenda. Journal of Environmental Psychology 29(3), 309–317 (2009)
5. de Vries, P., Aarts, H., Midden, C.J.: Changing Simple Energy-Related Consumer Behaviors How the Enactment of Intentions Is Thwarted. Environment and Behavior, 612–633 (2011)
6. Shiraishi, M., Washio, Y., Nakajima, T.: Using individual, social and economic persuasion techniques to reduce CO2 emissions in a family setting. In: Persuasive Technology 2009 (2009)
7. Foster, D., Lawson, S., Wardman, J., Blythe, M., Linehan, C.: Watts in it for me?: design implications for implementing effective energy interventions in organisations. In: CHI 2012 (2012)
8. Barendsen, E.: Saving Energy Where It Matters Most: A Hopeful Look, http://energy.gov/energysaver/articles/saving-energy-where-it-matters-most-hopeful-look-big-picture
9. Lucid Design Group: Elon University Strives to Reduce Energy Consumption Through Behavior Change (2010), http://www.luciddesigngroup.com/casestudies/elon.php
10. Winett, R.A., Neale, M.S., Grier, H.C.: Effects of self-monitoring and feedback on residential electricity consumption. Applied Behavior Analysis 12(2), 173–184 (1979)
11. Yun, R., Scupelli, P., Aziz, A., Loftness, V.: Sustainability in the Workplace: Nine Intervention Techniques for Behavior Change. In: Berkovsky, S., Freyne, J. (eds.) PERSUASIVE 2013. LNCS, vol. 7822, pp. 253–265. Springer, Heidelberg (2013)
12. Siero, F.W., Bakker, A.B., Dekker, G.B.: Changing organizational energy consumption behaviour through comparative feedback. Journal of Environmental Psychology 16, 235–246
13. Fischer, C.: Feedback on household electricity consumption. Energy Efficiency 1, 79–104 (2008)
14. Mercier, C., Moorefield, L.: Commercial Office Plug Load Savings Assessment (2011), http://www.efficientproducts.org/reports/plugload/PlugLoadSavingsAssessment.pdf
15. Fogg, B.J.: Persuasive technology: using computers to change what we think and do (2002)
16. Winett, R.A., Hatcher, J.W.: The effects of videotape modeling and daily feedback on residential electricity conservation. Applied Behavior Analysis 15, 381–402 (1982)
17. Kohlenberg, R., Phillips, T., Proctor, W.: A behavioral analysis of peaking in residential electrical-energy consumers. Journal of Applied Behavior Analysis 9(1), 13 (1976)
18. GreenBiz.com (2009), http://www.greenbiz.com/sites/default/files/GreenBizR
19. Harrigan, M.: Can we transform the market without transforming the customer? Home Energy 11(1) (1994)

Sustainability in the Workplace: Nine Intervention Techniques for Behavior Change

Ray Yun[1], Peter Scupelli[2], Azizan Aziz[1], and Vivian Loftness[1]

[1] School of Architecture,
[2] School of Design,
Intelligent Workplace, Carnegie Mellon University
{ryun,scupelli,azizan,loftness}@cmu.edu

Abstract. Human activity creates some of the greatest environmental challenges on the planet. Human-Computer Interaction (HCI) research on sustainability increased dramatically in the past five years. Researchers argue that technology plays a critical role in changing people's belief and behavior towards sustainability. Much of the sustainable HCI research currently focuses on domestic environments. However, 36% of electrical energy is used in the workplace and through office workers' behavior modification savings between 12% to 20% are possible. We investigate nine intervention techniques based on a review of studies in the fields in Persuasive technology, HCI, Ubicomp, environmental psychology, energy efficiency, and green building. We introduce related studies and design examples, and discuss design issues using our suggestions in designing persuasive systems for the workplace.

1 Introduction

The ASE (Alliance to Save Energy) report in 2009 [19] states that US office workers annually waste 2.8 billion dollars due to computers that are not shut off when they leave the office. For instance, a typical company with 10,000 computers wastes $260,000 and contributes 1.871 tons of CO_2 to climate change. People's habitual behavior wastes considerable resources and seriously damages the environment.

In the burgeoning Persuasive technology and sustainable HCI literature, many systems claim to motivate pro-environmental behaviors through innovative strategies. For example, ambient displays can inform users of energy use in real time [e.g. 25,50], and computer games can train teenagers about when and how they can save energy [e.g., 9]. However, DiSalvo et al. (2009) point out that much of the HCI literature provides limited information on evaluation in terms of the methods, number of participants, and the study length [18]. Furthermore, when HCI literature link design factors to behavior change, it rarely explains why such factors matter; notable exceptions include [10] and [12].

In environmental psychology, research on energy conservation behavior change connects theory and practice. Starting in the mid 1970s three persuasive behavior change techniques were developed and evaluated for environmental issues (e.g. frequent billing [37], rewards [13], pamphlets [59]). Given the historic context, many

S. Berkovsky and J. Freyne (Eds.): PERSUASIVE 2013, LNCS 7822, pp. 253–265, 2013.

environmental psychology studies were limited to traditional mediums (e.g., written advice [37] or educational video tapes [35]) and analogue interactions (e.g., paper bills delivered weekly or monthly [27]). The number of environmental behavior modification studies has significantly declined starting in the 1990s [27]. Researchers in the area of Persuasive technology and sustainable HCI [18, 26] recently argued that we can learn a lot from environmental psychology literature [e.g.,1], and introduced such literature to the HCI community [e.g., 26].

As Foster et al. [24] and Schwartz et al. [4] state, many existing studies in environmental psychology and HCI-related literature target domestic areas [e.g., 6,8,9,21,22,27,34,35,38,43,52]. Sustainability in the workplace setting is an area of HCI research that is not fully studied.

In the United States, commercial buildings account for 36% of all electricity used and are responsible for 18% of carbon emissions [20]. It is believed that office workers typically little care about saving energy because they are not responsible for paying energy-related bills [24]. Thus, in this paper, we discuss relevant intervention techniques and applications aimed at energy usage reduction for the workplace.

2 Past Sustainability Intervention Technique Reviews

We found reviews about intervention techniques to encourage pro-environmental behaviors in four bodies of literature: Persuasive technology [23,51], HCI [22,24,26], psychology [1, 60], energy efficiency [21,34]. In the field of energy efficiency, Wood et al. [34] introduced design elements for displaying energy information in six dimensions (the place of display, motivational factors, display units, display methods, timescale, category), and argued for effective techniques to save energy. In HCI, Fitzpatrick et al. [22] conducted an online survey of fifty-seven people and qualitative field studies (six households) on designing energy feedback displays, and discussed their findings in terms of four themes (behavior and context, data, device, engagement). Both studies focus on designing effective interventions that promote sustainable behavior but neither explains theoretically how their interventions cause desirable behavior.

Other studies attempt to link behavior change models. In environmental psychology, Abrahamse et al. [1] reviewed thirty-eight articles in environmental psychology and introduced six intervention techniques based on Skinner's behavior change model [46] - antecedent intervention (commitment, goal setting, information, modeling) and consequence intervention (feedback, reward). In the field of energy efficiency, Fischer [21] reviewed twenty-six papers in environmental psychology and detailed seven feedback types that can influence energy consumption reduction (frequency, duration, content, breakdown, presentation mode, comparison, and additional information). From a design perspective, both Abrahamse et al. [1] and Fischer [21] describe behavior change models [46,71] in limited detail. Fischer does not link the feedback elements to Matthies' model he introduced [55] and Abrahamse only uses Skinner's model [46] for categorizing aspects of the interventions.

In psychology, Kluger et al. [60] reviewed 131 studies and conducted meta-analysis on feedback-related interventions (correct solution, velocity, discouraging feedback, praise, verbal feedback, computer feedback, frequency, complexity, physical task, memory task, following rules, goal setting, thrust to self-esteem). In Persuasive

techonology, Fogg [23] introduced forty-two persuasive techniques that can affect people's behavior. To effectively apply Fogg's technologies in designing persuasive systems, Oinas-Kukkonen et al. [51] refined them into twenty-eight techniques with system requirements and examples in four categories (primary task support, dialog support, system credibility support, and social support). Those three studies cover various intervention techniques and related examples in general. Focusing on sustainability, Froehlich et al. [26] present HCI examples with six intervention techniques commonly used in environmental psychology (information, goal-setting, comparison, commitment, incentive, and feedback). Froehlich et al.'s research [26] provides a good link between HCI and environmental psychology, and introduces feedback-related interventions (eco-feedback technology) targeting typical individual users. To focus more on sustainability in the workplace, next we review the latest studies in Persuasive technology, environmental psychology, and HCI-related field, and suggest nine intervention techniques – education, advice, self-monitoring, goal setting, social-comparison, engagement, communication, control, and rewards. In the next section, we demonstrate the process we followed to obtain these intervention techniques.

3 Method

We conducted three iterative searches in our literature review. Our first search used keywords related to three different topics: the environment (e.g. energy conservation, green, sustainability), behavior change (e.g., persuasive, interventions), and workplace (e.g., office, employee). With the first search, we found four academic papers [24,29,30,102] and seven professional research reports [4,19,27,28,32,44,53]. We summarized their strategies to motivate office workers, and categorized them into nine intervention techniques, which can effectively be employed in office buildings.

Second, to understand what types of persuasive systems have been created with the intervention techniques, we reviewed articles in the field of Persuasive technology, HCI, and Ubicomp. Many papers we found focus on describing their design features, but not theoretically explaining why the design works that way, as Disalvo et al. mentioned in his paper [18]. This leads us to the next process.

Third, to theoretically explain the design factors for changing people's behavior, we searched literature in the field of environmental and social psychology and behavior science, and reviewed related articles and books. Doing this, we found various models that explain factors supporting behavior change, including Skinner's model that illustrates behavior occurrence factors (antecedent - behavior - consequence) [46], Stern's model that identifies seven barrier groups that prevent pro-environmental behaviors (household background, eternal constraints, worldviews, beliefs, knowledge, attention, resource-using behavior) [45], Piotrow's model that explains the process of health-behavior change (pre-knowledgeable, knowledgeable, approving, intending, practicing, advocating) [57], social cognitive theory that illustrates behavior change factors (personal factors, environmental factors) [15], the Transtheoretical model that explains a series of stages for behavior change (pre-contemplation, contemplation, preparation, action, maintenance, termination) [48], Fogg's Behavior Model (FBM) that presents three factors that lead to a target behavior (motivation, ability, and triggers) [41], and Geller's behavior change model that demonstrates four

performer stages (unconscious incompetence, conscious incompetence, conscious competence, unconscious competence) and three types of intervention (instructional, motivational, supportive) [31].

To understand one's environmental behavior change process, we took into account two issues: (a) Does the model clearly explain the stages of behavior change? (b) Is the model about sustainability? First, with regards to the stages of change, Stern's, Skinner's, Fogg's, and social cognitive theory models [15,41,45,46] focus on the relationship between the behavior and the behavior change factors. Instead, Piotrow's, Transtheoretical and Geller's models [57, 31, 48] show the process of behavior change. Second, regarding sustainability, the Piotrow's model and Transtheoretical model [57, 48] were developed for health behavior change; instead, Geller's model [31] focuses on environmental behavior change. Hence, we adapted Geller's model to explain our nine intervention techniques in the next section.

Next, we reviewed environmental psychology literature to identify the relative impacts of each factor towards energy saving. We noted what designers should consider for each technique. This review process was done iteratively and provided us with conceptual and practical knowledge in order to design persuasive systems for sustainability in the context of the workplace.

4 Intervention Techniques for Behavior Change

As mentioned previously, Geller's behavior change model apparently demonstrates the connection between the stages of change and types of interventions. In this section, we will adapt Geller's model to explain our nine intervention techniques. Geller calls to people that encounter an intervention "performers."

4.1 Behavior Change Model

Geller's behavior change model consists of four stages of performers and three types of intervention, which help performers move to a next stage (Fig. 1). First, unconscious incompetence is the stage where people do not behave in a sustainable way because they do not know how to do it. People in this stage can learn what they could do and/or why it is important. If they understand it, they move to the next performer stage, conscious incompetence. If not, they go back to the first stage. Even if people know what to do, they may not do it immediately because they need motivation. With the aid of motivational intervention, people can adopt pro-environmental behavior more easily and move to the next stage, conscious competence. In this stage, the person has now acquired environmental knowledge and performed a target behavior, but the goal of behavior change is to make this a habitual behavior. Supportive intervention can help people repeatedly perform the behavior and move to the goal stage, unconscious competence, once the pro-environmental behavior has become routine. There are three assumptions in Geller's model, the first is that people move linearly from one stage to the next, the second is that they do not relapse, and the third is that people fully understand desirable behavior as a binary state, yes or no. Based on the literature, these are contentious assumptions to make. People oscillate between understanding at a surface level to really understanding in depth, and they progress and regress. If someone

answers no to any question they can move back to any stage, not just to the previous one. Although the assumptions are unproved, Geller's model still simply demonstrates a performer's behavior change stages in the domain of sustainability. Of the components in the model, this paper focuses on the intervention techniques. Fig.1 shows nine techniques linked to the three primary interventions. We defined nine techniques and divided them into three groups based on their main characters and functions, but note that they are not mutually exclusive. In the following sections, we discuss the nine techniques more in detail in the context of the workplace.

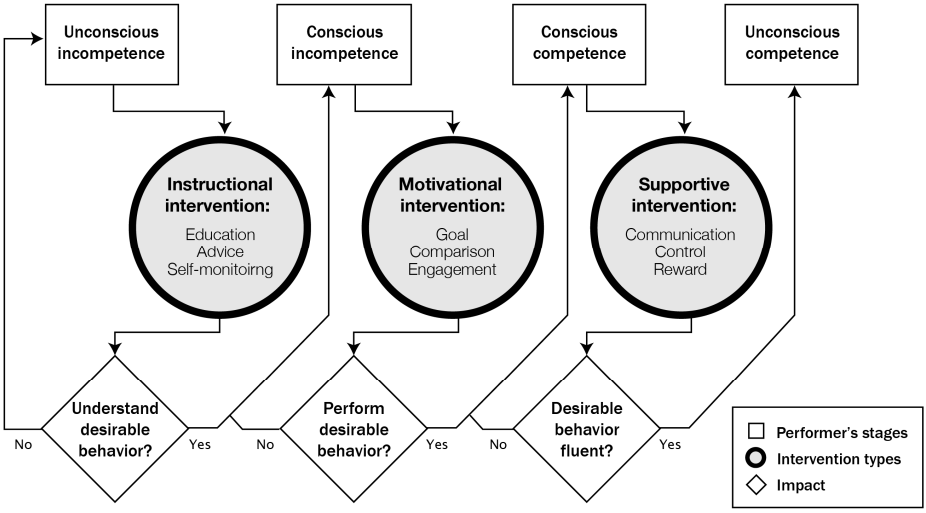


Fig. 1. Behavior change model for sustainability (simplified), adapted from Geller [31]. This demonstrates the four performer's stages and three types of interventions that help a performer move onto the next stage.

4.2 Instructional Intervention

This intervention helps people realize they have environmentally-destructive habits. The techniques for the intervention include education, advice and self-monitoring.

Education. This technique is an approach to teach people why we should conserve energy. According to Geller's model, education intervention can allow people with no background information about sustainability (unconscious incompetence) to understand why they should perform sustainable behaviors (conscious incompetence).

In environmental psychology, Winett et al. [58] studied video information that demonstrates that sustainable behaviors (e.g. turning down a thermostat) can reduce up to 28% of electricity consumption. Bandura et al [7] argued that demonstrating desired behavior works better than simply describing the behavior. Fogg [23] also argued in his book that if a system guides users through a process, it can have greater persuasive power.

Recent research [53] states many companies are not completely certain how to educate and engage employees to behave in a sustainable way. Forty-nine percent of

the 1300 survey respondents in the report said their companies have no or a not-advanced sustainability education program. Leading companies recently put significant effort into building websites that provide online training, run internal campaigns, and share information to improve companies' sustainability [32].

Advice. This strategy is an approach of giving people suggestions about what they can do to reduce energy consumption. According to the "Sustainability in the Workplace" report [44], one of the biggest barriers in the workplace that prevents pro-environmental behavior is a lack of 'how to' information (unconscious incompetence) Since office workers typically are not the ones who purchase and manage the appliances, written manuals about effective ways to use each appliance (e.g., how to setup the power management for their computers and monitors [28]) will be helpful for them to understand how to save energy (conscious incompetence). There are various methods to convey advice: emails, mobile instant messages, an agent's dialogue messages [3], icons on a mobile telephone [25], facial expressions [33], missions in a game [47] and many others.

Fischer [21] introduced Mosler and Gjutscher's study that shows how advice intervention can reduce electricity consumption. In this study, they provide participants with advice on how to reduce electricity usage. As a result, the group that received advice shows a difference of a 14.4% reduction from the control group without advice. To design effective advice intervention for the workplace, designers should think of two aspects: personalized advice and the timing of the advice. To design personalized advice, the system should know the user's context (e.g., what devices individual users have in the office) [2]. Harrigan, et al. [36] argues that providing individualized and personalized information can help motivate and guide people's energy consumption. Moreover, if the system knows what time they typically use a device, and what time they leave the office, the system can provide the advice at the appropriate time (e.g. sending a message to turn-off the device before the user leaves the office).

Self-monitoring. This allows people to observe their behavior performance, learn their energy consumption pattern, and explore where they could save energy (conscious incompetence). In environmental psychology studies, monitoring alone can contribute to some energy conservation. For example, Winett et al. [35] taught participants how to read electricity meters and asked them to read and plot the consumption daily. As a result, they showed a 7% reduction compared with the control group.

The literature argues self-monitoring can be more powerful and lead to bigger energy savings if it provides real-time information, shows appliance-specific data, and compares current with historic data. For example, energy efficiency-related studies [21, 22, 23] argue that real-time feedback improves interaction between the user and the system [27] and helps motivate the user. Self-monitoring with appliance-specific data is also useful because it allows people to see how or where they could improve their energy conservation [21]. Historic data can be helpful as well because users can better learn about their consumption pattern. Wilhite et al. [37] argues that providing a historic comparison with last year's data reduced energy consumption by 4%.

According to Foster's research [24], office workers would prefer a monetary value for monitoring their energy consumption, because it may be more useful and makes the resource more tangible. Various units such as dollars, CO_2, or hamburgers can be used depending on who the users are [54].

4.3 Motivational Intervention

Even if people realize their habitual behaviors can harm the environment and know what to do via instructive intervention, they may not do it. In this case motivational intervention is needed to stimulate people to make pro-environmental changes. (It turns "conscious incompetence into conscious competence.") The specific techniques for this type of intervention are goal-setting, comparison and engagement.

Goals. This intervention is a strategy to motivate people by challenging them to achieve their aim. This strategy has been studied in psychology for a relatively long time. For example, Van Houwelingen et al. [38] assigned a 10% reduction goal to one group of participants and no goal to another group, and showed a 12.3% energy savings difference (natural gas) is made compared with the control group. McCalley et al. [17] studied possibility of using goal setting and an additional intervention, foot-in-the-door technique (FITD- giving a small request first and larger request later). They found out FITD encourages people to set higher goals than usual but is not effective for energy conservation (water) when the two techniques are used together. They suggest employing FITD where goal intervention is not used.

According to Foster et al. [24], many office workers feel frustrated when the company assigns unrealistic or unachievable sustainability goals. Shiraishi et al. [47] argues that a too-challenging and long-term goal may make people unmotivated. To improve it, companies can quantify and visualize each employee's performance and assign short-term personal goals to individuals so that they can see their achievement in real-time [53].

Comparison. This is a strategy to show people their and others performance and motivate people to change their behavior. Foster et al [24] showed office workers could be effectively motivated to perform pro-environmental behavior when they realize what others do for sustainability (conscious incompetence into conscious competence).

Siero et al. [30] show the comparison intervention can help motivate employees to conserve energy. They provide different information to two units in a metallurgical company. Employees in the first unit received information about their energy consumption and assigned goal and the other unit additionally received information about the performance of other units. As a result, the second unit who received comparative information achieved more energy savings.

To use comparison strategy in the workplace, comparison targets should be carefully selected. For example, people with similar a jobs and work environments can make relatively fair comparisons. If compared subjects are not similar, e.g., divisions that have different numbers of people, comparisons can be made with converted units such as energy consumption reduction rate per month, or average consumption per person.

Engagement. This is a strategy to appeal to people's emotion or curiosity, and motivate them to perform a pro-environmental behavior. Engagement has not been considered as a primary intervention for motivating energy conservation in environmental psychology. In the Persuasive, HCI and Ubicomp community, however, it has been commonly used as a motivational factor by many researchers and designers. The genres are various: Interactive agent/avatar [e.g. 3,6,33], ambient display [e.g. 25,50],

eco-art or eco-visualization [e.g. 42], game [e.g. 8,9,47] and many others. With an engaging manner, the system can naturally influence behavioral change. For example, virtual characters (e.g. a polar bear [25], coral [50], or human [33]) can make users feel empathy and perform desired behaviors to keep the character alive or happy. Watkins et al [39] pointed out that visual attractiveness is a powerful factor for persuasion and Bang argued [9] that game characters and agents can be more persuasive if they look visually appealing. A recent report [32] states that an eco-art installation in the workplace (which contains a message about sustainability) can easily draw employees' attention, help raise their awareness of sustainability and motivate them to perform desirable behavior (unconscious incompetence into conscious incompetence and conscious competence via engagement intervention).

As stated above, various approaches have been used, however, it is still questionable whether engagement can motivate users for a long time. This intervention alone may be insufficient for long-term behavior change, so one way designers can think of it is to harness engagement with supportive intervention so that users can be motivated continuously.

4.4 Supportive Intervention

Motivational factors have been shown to be effective in supporting pro-environmental behavior, but are often temporary. However, the desired results are to sustain these new behaviors for the long-term. The techniques for the intervention consist of communication, control and reward.

Communication. This strategy is to conserve energy in the workplace by providing a communication tool with people with same interest (social network) or people who have the authority to change facilities. Social network systems are now common and allow people to easily share information, comments and testimonies. Froehlich et al. [26] argues that social network systems (e.g. Facebook, Twitter) are a good place for social sharing and public commitment about sustainable behavior. Mankoff et al. [56] argue social networks can increase continuous participation in a social movement. Social networks in the workplace can be beneficial to motivate people in their daily life and can be a good supportive tool to encourage each other in the community for the long term (conscious competence into unconscious competence). To harness this strategy for the workplace, however, designers should be aware of the business' culture because the use of social media during the work hours may be allowed in some organizations (e.g. universities) but not in others (e.g. typical offices in companies) [24].

Additionally, methods that allow communication with people who manage facilities in the workplace may increase sustainability. Typically people who are in charge of devices in workplace are not office workers but people in the facility management department. When office workers self-monitor their plug loads and realize a certain item consumes too much energy, they can report this and ask their facility manager to fix it or replace it with a more energy-efficient product. Oinas-Kukkonen [51] also pointed out if a system plays a social role, users will more likely use it for persuasive purposes.

Control. This technique is an approach to conserving energy by providing easy and simple ways to control their energy consumption. For example, a system interface can provide a button that controls users' devices based on their settings, and they can press

a button to turn on or off a group of devices at once when they come into or leave the office. This transforms the target behaviors into ones that can be achieved more easily and simply, and users can be motivated to do it. This is similar to Fogg's reduction principle [23]. Reducing complex behavior into simple tasks increases the cost/benefit ratio of the behavior and motivates users to perform it (conscious competence) and to repeat it to get the benefit continuously. Although this approach has huge potential for motivating sustainable behaviors, it has not been thoroughly studied yet.

Mercier et al. [4] demonstrate that advanced controls for office appliances such as remote control plug strips or timer plug strips can reduce respectively up to 55% and 43% of electricity consumption in the office building.

To employ this strategy in the persuasive system, designers should consider the technical challenges of making everything controllable by a single system. However, at least plug loads (electricity) can be easily controlled with many commercial products such as Plugwise (http://plugwise.com) or Enmetric (http://enmetric.com). These are small devices that can be connected between an outlet and an appliance's plug, and give users not only control but also real-time consumption data so that people can easily monitor and control their plug loads. We expect that soon there will be controls for other energy resources such as natural gas and water.

Reward. A reward is a prize for the performance of target behavior. Unlike other elements that are designed for "before behavior," this element occurs "after behavior." B. F Skinner [46] categorized it as consequence intervention. Because this intervention results from the performer's behavior and the performer gets motivated by it to perform the behavior again, we categorize it as a supportive intervention.

Foster's study [24] shows that office workers think about what the potential benefit will be when they are asked to do a sustainable behavior. Effective rewards that they suggested included saving funds for a Christmas party or receiving coupons for free food in the cafeteria, both relatively realistic and tangible.

Environmental psychology studies show the reward intervention can motivate people to conserve energy. Kohlenberg et al. [14] showed that monetary payment can reduce energy usage. They gave different rewards according to people's energy savings. (e.g., 5 -10%: $2/week, 10-15%: $3/week, 15-20%: $5/week). They found that the reward group showed a reduction of 15% compared with the control group.

Several studies [45, 37] also show that the larger the reward offered, the better people perform pro-environmental behavior, and a low incentive does not affect people's behavior. In Persuasive and HCI, virtual rewards [47], and emotional rewards [25,50] were introduced but there are not many studies on how virtual or emotional rewards can motivate people and actually contribute to energy conservation.

5 Discussion

In this section we will discuss the important issues that are related to the intervention techniques in general but not covered in the previous section.

First, the leadership role is the most important factor for sustainability in the workplace [24]. Most of the interventions we discussed in the previous section can hardly be used without the leadership's permission because the office workers primary duty is not saving energy, but doing their job. For example, people might want to

open a web browser to self-monitor their performance in real-time, get feedback from the system, or share comments on each other's performance. If their leadership does not allow green activities, it will be hard for employees to perform them. Therefore, it is important that the leadership is conscious about environmental issues and establishes eco-policies to support the environmental strategies we are suggesting for their organization's sustainability.

Second, we did not rank-order the intervention techniques by which interventions can save more energy in a certain situation, because there exist few studies in the workplace domain [4,24]. However, in the context of a residential building, Ehrhardt-Martinez et al. [27] conducted a meta analysis on thirty-six studies and provided comparative statistics on the five types of feedback - *enhanced billing* (comparative, historic data), *estimated feedback* (disaggregated by appliances), *daily/weekly feedback* (daily/weekly advice), *real-time feedback* (real-time, not disaggregated), *real-time plus* (real-time, disaggregated by appliances). Respectively they contribute to 3.8%, 6.8%, 8.4%, 9.2%, 12.0% energy savings. Though this study focuses only on the feedback-related interventions (self-monitoring, advice, comparison) and the context of household users, it helps to explain which persuasive factors can potentially motivate people to save more energy.

Third, habits can thwart environmental behavior change. Bamberg et al. [61], Matthies et al. [16], and de Vries et al. [43] argued even if people have intention, belief, or plan to change their behavior, they often don't change it when it is their habit (e.g., travel mode change [16,61], switching the light off [43]). For those who already have bad environmental habits, rather than focusing on changing their intentions and attitude via informational interventions (e.g., education, advice, self-monitoring, comparison), changing behavior cost and circumstance may work better via structural interventions [52] (e.g., reward for best energy saving contribution, or one-click control for all my electronics).

6 Conclusion

Individuals can contribute significantly to energy conservation by changing their behavior environmentally in the workplace. In this paper, we investigated intervention techniques that can promote sustainable behaviors, listed design issues and made suggestions for each intervention technique. The challenge still remains - understanding what interventions or combinations of interventions are most effective for certain type of workers. To follow up on this study, we have built a user-based energy dashboard applying the above intervention techniques and are currently conducting a series of field studies to evaluate how much energy can be saved by the intervention techniques [62]. These field studies also assess occupants' perceived comfort while they save energy in the workplace, through IEQ (indoor environmental quality) surveys in the domain of lighting, temperature, and ventilation.

Acknowledgments. This study is supported by the Greater Philadelphia Innovation Cluster for Energy Efficient Buildings project (GPIC, a U.S. DOE Innovation Hub, Subtask 6.4) under U.S. Department of Energy Award Number EE0004261.

References

1. Abrahamse, W., Steg, L., Vlek, C., Rothengatter, T.: A review of intervention studies aimed at household energy conservation. Environmental Psychology 25(3), 273–291 (2005)
2. Adomavicius, G., Tuzhilin, A.: Context-Aware Recommender Systems. In: Ricci, F., Rokach, L., Shapira, B., Kantor, P.B. (eds.) Recommender Systems Handbook, pp. 217–253. Springer US (2011)
3. Al Mahmud, A., Dadlani, P., Mubin, O., Shahid, S., Midden, C., Moran, O.: iParrot: Towards Designing a Persuasive Agent for Energy Conservation. In: de Kort, Y.A.W., IJsselsteijn, W.A., Midden, C., Eggen, B., Fogg, B.J. (eds.) PERSUASIVE 2007. LNCS, vol. 4744, pp. 64–67. Springer, Heidelberg (2007)
4. Mercier, C., Moorefield, L.: Commercial Office Plug Load Savings Assessment (2011), http://www.efficientproducts.org/reports/plugload/PlugLoadSavingsAssessment.pdf
5. Schwartz, T., Betz, M., Ramirez, L., Stevens, G.: Sustainable energy practices at work: understanding the role of workers in energy conservation. In: NordiCHI 2010 (2010)
6. Ruijten, P.A.M., de Kort, Y.A.W., Kosnar, P.: Bridging the Gap between the Home and the Lab: A Qualitative Study of Acceptance of an Avatar Feedback System. In: Bang, M., Ragnemalm, E.L. (eds.) PERSUASIVE 2012. LNCS, vol. 7284, pp. 251–255. Springer, Heidelberg (2012)
7. Bandura, A.: Self-efficacy: Toward a unifying theory of behavioral change. Psychological Review 84, 191–215 (1977)
8. Gamberini, L., Spagnolli, A., Corradi, N., Jacucci, G., Tusa, G., Mikkola, T., Zamboni, L., Hoggan, E.: Tailoring Feedback to Users' Actions in a Persuasive Game for Household Electricity Conservation. In: Bang, M., Ragnemalm, E.L. (eds.) PERSUASIVE 2012. LNCS, vol. 7284, pp. 100–111. Springer, Heidelberg (2012)
9. Bang, M., Gustafsson, A., Katzeff, C.: Promoting New Patterns in Household Energy Consumption with Pervasive Learning Games. In: de Kort, Y.A.W., IJsselsteijn, W.A., Midden, C., Eggen, B., Fogg, B.J. (eds.) PERSUASIVE 2007. LNCS, vol. 4744, pp. 55–63. Springer, Heidelberg (2007)
10. Thieme, A., Comber, R., Miebach, J., Olivier, P.: We've bin watching you: designing for reflection and social persuasion to promote sustainable lifestyles. In: CHI 2012, pp. 2337–2346 (2012)
11. Berry, L.: The market penetration of energy-efficiency programs (No. ORNL/CON-299). Oak Ridge National Lab., TN (USA) (1990)
12. 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 (2010)
13. Seaver, W.B., Patterson, A.H.: Decreasing fuel-oil consumption through feedback and social commendation. Journal of Applied Behavior Analysis 9(2), 147 (1976)
14. Kohlenberg, R., Phillips, T., Proctor, W.: A behavioral analysis of peaking in residential electrical-energy consumers. Journal of Applied Behavior Analysis 9(1), 13 (1976)
15. Pajares, F.: Overview of Social Cognitive Theory and of Self-efficacy (2002)
16. Matthies, E., Kuhn, S., Klockner, C.A.: Travel Mode Choice of Women: The Result of Limitation, Ecological Norm, or Weak Habit? Environment and Behavior 34(2), 163–177 (2002)
17. McCalley, L.T.: Consumer Response to Product-Integrated Energy Feedback: Behavior, Goal Level Shifts, and Energy Conservation. Environment and Behavior 43(4), 525–545 (2011)
18. DiSalvo, C., Sengers, P., Brynjarsd'ottir, H.: Mapping the landscape of sustainable HCI. In: CHI 2010, pp. 1975–1984. ACM, New York (2010)

19. 1E PC Energy Report: PC Energy Report 2009, USA, United Kingdom, Germany (2009), http://www.1e.com/energycampaign/downloads/PC_EnergyReport2009-US.pdf
20. Barendsen, E.: Saving Energy Where It Matters Most: A Hopeful Look (2009), http://energy.gov/energysaver/articles/saving-energy-where-it-matters-most-hopeful-look-big-picture
21. Fischer, C.: Feedback on household electricity consumption: a tool for saving energy? (2008)
22. Fitzpatrick, G., Smith, G.: Technology-Enabled Feedback on Domestic Energy Consumption: Articulating a Set of Design Concerns. Pervasive Computing 8(1), 37–44 (2009)
23. Fogg, B.J.: Persuasive technology: using computers to change what we think and do. Ubiquity (December 2002), doi:10.1145/763955.763957
24. Foster, D., Lawson, S., Wardman, J., Blythe, M., Linehan, C.: "Watts in it for me?": design implications for implementing effective energy interventions in organisations. In: CHI 2012, New York, NY, USA, pp. 2357–2366 (2012)
25. 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)
26. Froehlich, J., Findlater, L., Landay, J.: The design of eco-feedback technology. In: CHI 2010, pp. 1999–2008. ACM, New York (2010), doi:10.1145/1753326.1753629
27. Ehrhardt-Martinez, K., Donnelly, K.A., Laitner, J.A.S.: Advanced Metering Initiatives and Residential Feedback Programs, Washington, D.C. (2010)
28. Mercier, C.: Commercial office plug load savings and assessment: Final report (July 2011)
29. Lehrer, D., Vasudev, J.: Evaluating a social media application for sustainability in the workplace. In: CHI EA 2011, pp. 2161–2166. ACM, New York (2011)
30. Siero, F.W., Bakker, A.B.: Changing organizational energy consumption behaviour through comparative feedback. Journal of Environmental Psychology 16, 235–246 (1996)
31. Geller, E.S.: The challenge of increasing proenvironmental behavior. In: Betchel, R.B., Churchman, A. (eds.) Handbook of Environmental Psychology, New York, pp. 525–540 (2002)
32. GreenBiz.com: Green Teams – Engaging Employees in Sustainability (2009), http://www.greenbiz.com/sites/default/files/GreenBizR
33. Yun, R., Gross, M.D.: RayMatic: ambient meter display with facial expression and gesture. In: CHI EA 2011, pp. 341–346. ACM, New York (2011)
34. Wood, G.: Energy-use information transfer for intelligent homes: Enabling energy conservation with central and local displays. Energy and Buildings 39(4), 495–503 (2007)
35. Winett, R.A., Neale, M.S., Grier, H.C.: Effects of self-monitoring and feedback on residential electricity consumption. Journal of Applied Behavior Analysis 12(2), 173–184 (1979)
36. Harrigan, M.: Can we transform the market without transforming the customer? Home Energy 11(1) (1994)
37. Wilhite, H. (n.d.): Advances in the use of consumption feedback information in energy billing: the experiences of a Norwegian energy utility — ECEEE. Proceeding
38. van Houwelingen, J.H., van Raaij, W.F.: The effect of goal-setting and daily electronic feedback on in-home energy use. Journal of Consumer Research 16 (1989)
39. Watkins, L.M.: Screening Job Applicants: The Impact of Physical Attractiveness and Application Quality. International Journal of Selection and Assessment 8(2), 76–84 (2000)
40. Hayes, S.C., Cone, J.D.: Reducing residential electrical energy use: payments, information, and feedback. Journal of Applied Behavior Analysis 10(3), 425–435 (1977)
41. Fogg, B.J.: A behavior model for persuasive design. In: Proceedings of the 4th International Conference on Persuasive Technology, p. 40. ACM (April 2009)

42. Holmes, T.G.: Eco-visualization: combining art and technology to reduce energy consumption. In: C&C 2007, pp. 153–162. ACM, New York (2007)
43. de Vries, P., Aarts, H., Midden, C.J.: Changing Simple Energy-Related Consumer Behaviors How the Enactment of Intentions Is Thwarted by Acting and Non-Acting Habits. Environment and Behavior 43(5), 612–633 (2011)
44. Sustainability in the workplace report (2011), http://www.sustainabilityatwork.com.au/downloads/research-report.pdf
45. Stern, P.C., Oskamp, S.: Managing scarce environmental resources. In: Stokols, D., Altman, I. (eds.) Handbook of Environmental Psychology, New York, pp. 1043–1088 (1987)
46. Skinner, B.F.: Science And Human Behavior. Simon and Schuster (1965)
47. Shiraishi, M., Washio, Y., Takayama, C., Lehdonvirta, V., Kimura, H., Nakajima, T.: Using individual, social and economic persuasion techniques to reduce CO2 emissions in a family setting. In: Persuasive, pp. 13:1–13:8. ACM, New York (2009)
48. Prochaska, J.O., Velicer, W.F.: The transtheoretical model of health behavior change. American Journal of Health Promotion 12(1), 38–48 (1997)
49. Kimura, H., Nakajima, T.: Designing Persuasive Applications to Motivate Sustainable Behavior in Collectivist Cultures. The Other Side of Technology 9(1), 7–28 (2011)
50. Kim, T., Hong, H., Magerko, B.: Designing for Persuasion: Toward Ambient Eco-Visualization for Awareness. In: Ploug, T., Hasle, P., Oinas-Kukkonen, H. (eds.) PERSUASIVE 2010. LNCS, vol. 6137, pp. 106–116. Springer, Heidelberg (2010)
51. 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)
52. Steg, L., Vlek, C.: Encouraging pro-environmental behaviour: An integrative review and research agenda. Journal of Environmental Psychology 29(3), 309–317 (2009)
53. NEEFUSA (2009), http://www.neefusa.org/BusinessEnv/EngagedOrganiZation_03182009.pdf
54. Lucid, Building Dashboard site, http://www.buildingdashboard.com/clients/emory/turman/
55. Matthies, E.: How can psychologists improve their outreach towards practitioners? A Suggestion for a Model of Environmentally Sound Everyday Practice 9(1), 62–81 (2005)
56. Mankoff, J., Matthews, D., Fussell, S.R., Johnson, M.: Leveraging Social Networks To Motivate Individuals to Reduce their Ecological Footprints. In: HICSS 2007 (2007)
57. Piotrow, P.T., Kincaid, D.L., Rimon II, J.G., Rinehart, W.: Health communication: lessons from family planning and reproductive health (1997)
58. Winett, R.A., Fort, T.R.: The effects of videotape modeling and daily feedback on residential electricity conservation. Applied Behavior Analysis 15, 381–402 (1982)
59. Heberlein, T.A.: Conversation information: The energy crisis and electricity consumption in an apartment complex. Energy Systems and Policy 1, 105–117 (1975)
60. Kluger, A.N., DeNisi, A.: The effects of feedback interventions on performance: A historical review, a meta-analysis, and a preliminary feedback intervention theory. Psychological Bulletin 119(2), 254 (1996)
61. Bamberg, S.: Is a Residential Relocation a Good Opportunity to Change People's Travel Behavior? Environment and Behavior 38(6), 820–840 (2006)
62. Yun, R., Lasternas, B., Aziz, A., Loftness, V., Scupelli, P., Rowe, A., Kothari, R., Marion, F., Zhao, J.: Toward the Design of a Dashboard to Promote Environmentally Sustainable Behavior among Office Workers. In: Berkovsky, S., Freyne, J. (eds.) PERSUASIVE 2013. LNCS, vol. 7822, pp. 246–252. Springer, Heidelberg (2013)

Author Index