The following full text is a publisher’s version.

For additional information about this publication click this link.
http://hdl.handle.net/2066/161893

Please be advised that this information was generated on 2019-04-01 and may be subject to change.
Adaptivity and Personalization in Persuasive Technologies

Shlomo Berkovsky, Maurits Kaptein†, Massimo Zancanaro

† Statistics and Research Methods, Tilburg University, Tilburg, the Netherlands
m.c.kaptein@uvt.nl

Abstract. Persuasive technologies are used to persuade people to maintain a healthier lifestyle, purchase certain items, keep a sustainable environment, and more. However, persuasive interventions are generally delivered in a non-personalized, one-size-fits-all manner, which may limit their effectiveness. We argue that the application of personalization in persuasive technologies has the potential to substantially boost their impact. To this end, this paper defines a range of dimensions that need to be considered by designers of persuasive systems and analyzes three works that exploit personalization for persuasion purposes.

1 Introduction

Many online services apply persuasion to a range of purposes, in order to influence the behavior or attitudes of their users. Although persuasive technologies are successful in influencing user behavior, we argue that this influence could be amplified if the persuasive interventions were applied in a personalized manner [2,10]. That is, the type, intensity, communication, and presentation of the persuasive interventions can and should be adapted to characteristics of an individual user. For example, if an eCommerce site was able to understand a user’s susceptibility to various forms of persuasion, it could apply appropriate language, avatars, or arguments to persuade the user to purchase a product. Similarly, a system encouraging healthy lifestyle could achieve increased impact, if it was aware of the most influential persuasive strategy for every user.

With increased regularity, we see personalization applied to persuasive systems. Personalized feedback motivating users to quit smoking was found to lead to higher cessation levels than general information [4]. Models for a personalized persuasive argument, considering the user’s involvement with the topic, were developed in [11]. Mobile text messages discouraging snacking, tailored to a user’s susceptibility to persuasive strategies, were found more effective than non-personalized ones [8]. Personalized rewards were applied in physical activity motivating games, to persuade children to perform mild activity while playing [1]. Advertisements tailored to the user’s personality traits were evaluated positively, when they cohered with user’s motives [7]. Persuasive strategies were applied in serious games and tailored to gamer’s personality, to stimulate target behavior of every gamer [12]. And there are more examples.
Following an analysis of these works, we present a list of principal dimensions (or design considerations) that shape the application of personalization in persuasive technologies. We initially introduce these dimensions, and revisit them in the following sections, when discussing three specific works. We split them into three categories of dimensions: (a) what user data serves as the basis for personalization?; (b) how are the interventions of a persuasive system personalized?; and (c) how are the personalized persuasive interventions evaluated?

What facilitates personalization? A variety of traits and personal characteristics can be encapsulated in the user models and exploited for personalization purposes. Here, designers of persuasive systems need to answer a number of questions: What features will drive personalization and should be modeled by the system? Is this information gathered explicitly or implicitly? How dynamic are the user models? Answers to these questions practically derive the type of personalization offered by persuasive technologies.

How is personalization achieved? The persuasive interventions can be personalized in a various ways, and using a broad spectrum of techniques. What is the granularity of personalization that is envisioned? Is this an end- or means- persuasion? What term of behavior change (long or short) does the persuasion aim to achieve? Is the persuasion direct or indirect? Does the persuasive system disclose that the interventions are personalized? All these decisions can affect the exact way the persuasion gets personalized.

How is personalization evaluated? Various methodologies can be chosen in order to evaluate the effectiveness of the personalized persuasion. Is the effectiveness of persuasive interventions evaluated subjectively or objectively? What are the success criteria and the exact evaluation metrics exploited? What is the considered term of evaluation? These factors impact both the scientific and practical validity of the obtained findings.

In this paper we introduce three published case studies of personalized persuasive technologies [1,8,15], along with their mapping onto the above dimensions. Through this mapping, we exemplify how various design choices can be operationalized, and highlight the gaps in the current research and practical application of personalized persuasive technologies.

2 Case Study 1: Physical Activity Motivating Games

Many information technology solutions helping to combat the obesity problem have been developed. For example, NEAT-o-Game is an active game, in which player’s activity is captured by an accelerometer and affects the speed of the game character [6]. GeoKaos and Flareqoor are two games that exploit physiological sensors to adjust the intensity of exercising [3]. Activity motivating games were designed to leverage the playability of
games to motivate players to perform mild physical activity while playing [1]. This motivation was achieved by offering virtual rewards for real physical activity performed by players.

This latter idea was applied to an open source game, which offered to players time-based rewards, i.e., players could gain extra time in return for jumping. The efficacy of the activity motivating games was evaluated thorough the amount of activity performed and the enjoyment of playing. The evaluation showed that players performed more activity and did not report a decrease in the enjoyment of playing. However, further analysis raised two concerns: (1) experienced players performed less activity than novice players, and (2) the highest enjoyment was obtained when the game difficulty was adapted to players’ gaming skills. These triggered further work on personalization of the activity motivating games.

2.1 Implementation and Evaluation of Personalization

To achieve these goals, we introduced two personalized techniques: Tailored Reward (TR) and Personalized Difficulty (PD). The goal of TR was to balance the amount of activity performed by various players. In context of time rewards, this was achieved by modifying the reward time gained for each jump, i.e., reducing the reward time of experienced players requiring them to perform more activity. TR predicted the reward times using a stereotype-based algorithm that classified players into three skill clusters and assigned the rewards time accordingly. Note that TR was not truly personalized, but only tailored to the player’s cluster. The goal of PD was to set the game level difficulty, i.e., level time limit, in a player-dependent manner. PD predicted the personalized level completion time for a player using the K-Nearest-Neighbors method: computed player similarity, selected most similar players, and aggregated their level completion times. Here, the time limit was set in a personalized way, as the completion time was predicted for every player individually. 90 players participated in the evaluation and were divided into two groups: TR with tailored rewards and PD with personalized time limits. The baseline group BL had fixed rewards and time limits.

To evaluate the impact of TR, we computed the average number of jumps performed in each skills clusters (see Figure 1). The overall trend was still that higher-skilled players performed less activity than lower-skilled players. However, the TR technique substantially reduced the differences between the clusters. The number of jumps decreased by 11% in the low cluster, increased by 13% in the medium cluster, and increased by 29% in the high cluster. Hence, tailoring the reward time was an important step towards motivating all players to perform a comparable amount of activity. Although the PD technique caused more players to gain time through physical activity, the change in the enjoyment of playing was not significant. On the one hand, personalizing the level time limits required players to jump and interrupted the flow of playing, which could have decreased the enjoyment. On the other hand, this also personalized player interaction with
the game, which could have increased the enjoyment. All in all, these two factors outweighed each other, such that applying PD retained the enjoyment of playing.

2.2 Discussion of the Personalization Dimensions

We now revisit the dimensions of personalization outlined in the introduction, and discuss their implementation in the activity motivating games case study.

![Figure 1. Number of jumps observed in the gaming skill clusters.](image)

What facilitates personalization? In this case study, the reward times and levels time limits were based on the observed gaming skills, which were captured implicitly through the completion times of Neverball levels. Although additional information about the players was collected in the pre-study questionnaire, this information had low correlation with the measured gaming skills. Note also that the model captured through the observed level completion times was gradually refined with every new level completed. These ongoing updates of the model ensured its freshness and accuracy.

How is personalization achieved? The adaptivity of activity motivating games was both tailored and personalized; TR applied cluster-based tailoring, whereas PD computed personalized level time limits. This case study used means-persuasion, as the target action was always jumping, but the extra-time gained for jumps and the time limit triggering the jumps were manipulated. The personalization was not disclosed to players, as they were not aware of these manipulations. The persuasion was indirect, since the target physical activity action was masked behind the primary action of playing. The impact of persuasion was only short-term, as the players were motivated to be active only while playing, whereas the motivation disappeared beyond this context.

How is personalization evaluated? The impact of personalization in activity motivating games was evaluated objectively, through the number of jumps performed by players. There also was subjective evaluation of the perceived enjoyment, but this did not
directly relate to the persuasive interventions. The term of the evaluation was short, as the impact of activity motivating games on the players’ lifestyle and daily routine was not measured.

3 Case Study 2: Implicit Direct-Means Personalization

There is a broad literature on persuasion principles as distinct means towards the same end: they provide different ways, in which persuasive requests for the same goal can be framed. This case study focuses on the use of three persuasion principles: scarcity (people value things that are scarce), authority (people tend to comply with authorities), and consensus (people tend to do as others do).

It is well-known that users differ in their responsiveness to these persuasive principles [9]. This finding motivates adapting the use of persuasion principles to individual users via personal profile that describes which influence principles are effective for each user. The ways to create persuasion profile, in order to capture these individual differences, were presented in [8]. The resulting profile can be used to select a persuasion principle for every user. This case study examines the effects of building persuasion profiles based on the observed responses of users to messages, and uses these profiles to inform subsequent message selection.

3.1 Implementation and Evaluation of Personalization

The case study examined the responses of 1129 users of a health service to reminder emails they received to upload data from their so-called activity monitors. The messages were structured as follows:

Dear [name]. How are you doing? We hope all is well. It is 3 days since the last time you connected your Activity Monitor.

[Persuasive paragraph]

We would like to remind you to connect it to your PC and stay in touch.

Implementation of the persuasion principles was inserted as the [persuasive paragraph]. These texts were “We would like to remind you to connect your Activity Monitor to your PC soon. Today is a great day to stay fit so make sure you do not miss out on your participation!” for scarcity, “Experienced coaches recommend frequent uploads of your activity data. This will help you to gain more insight and be more active!” for authority, and “People like you who connect their Activity Monitor frequently with their PC are more likely to benefit from the program and obtain a healthy lifestyle!” for consensus. Multiple messages per principle were used for stimulus sampling. After sending the email, user behavior was monitored and the success, i.e., uploading within 24 hours after opening the message, was measured.

To evaluate the effect of personalized persuasion, the users were randomly allocated into four conditions:
1. Control: Standard reminders not containing any persuasion principles.
2. Best pre-tested: One of the messages implementing the authority principle, which was judged most motivating in a pre-study pilot.
3. Random: Randomly selected message implementing one of the persuasion principle with probabilities equal for all the principles.
4. Adaptive: Messages tailored to the user’s persuasion profile, which was dynamically updated after each user interaction.

To test the effects of personalization, the differences in success rates across the conditions were compared. Figure 2 shows the estimated success probability as a function of the number of messages sent to users. The adaptive messages were more successful than others and they remained successful for longer. The decrease in the adaptive condition was significantly slower than in the control, best pre-tested, and random conditions, indicating that personalized reminders were more effective than the non-personalized ones.

Fig. 2. Overview of the success rates of the reminder messages.
3.2 Discussion of the Personalization Dimensions

What facilitates personalization? In this case study the modeling of personality was based directly on user responses. This is somewhat atypical, as personality traits were not measured using existing scales, but trait differences were deduced from user the observed responses to authority messages. Thus, the modeling of user traits was implicit and the users were largely unaware of the modeling. The case study also used dynamic modeling of user traits – the user profile was updated after each interaction with the user, such that the profiling was not a one-off activity.

How is personalization achieved? The personalization was achieved by sending distinct messages to the users. The granularity of the personalization was quite low: the only dimension that was personalized was the content of the message, while other factors like channel and timing were not personalized. This case study provides a strong example of means-based personalization: the end goal of every message was identical, but the phrasing of the messages differed across the users. Contrary to the other case studies discussed in this paper, the impact term of this case study was very short: the messages only attempted to change one instance of a simple behavior. In this case, direct persuasion was used, as the message stated directly the intended behavior. However, the personalization was not disclosed to the users.

How is personalization evaluated? The evaluation of this case study was objective, since direct measures of the intended behavioral change were used as the primary outcome. The metric was the upload of the activity data by the users and the value of personalization was assessed through the success of messages over time. The term of the evaluation was about six months, although the effective term was determined by the number of messages received by users, which varied from 3 to 24. Thus, the evaluation was long-term, although this was primarily driven by the need for building implicit profiles; the long-term evaluation was not set up for measuring long-term behavior change effects.

4 Case Study 3: The Augmented Café Table

An ambient display is an information system that displays non-critical information by providing non-obtrusive visualizations to users. Ambient information should preferably be implicit rather than explicit, in order to avoid distressing the user with excessive attentional demands. In this sense, ambient displays are usually referred as peripheral displays, since they are meant to provide information in the periphery of the visual attention of the users.

The use of ambient displays for influencing user behavior – suggesting to take stairs rather than elevator – using ambient lights was studied in [13]. More closely related work investigated the use of peripheral displays that visualize the amount of participation in group discussions, with the explicit objective of balancing the contributions of the
participants [5]. The work described in this case study has a similar goal: to influence group behavior by an ambient display that visualizes implicit information. However, it aims at suggesting topics for conversation using information about the group.

4.1 Implementation and Evaluation of Personalization

The study was set up in a museum setting. The persuasive system, called the Augmented Café Table, used the table as a peripheral display that visualized stimuli, in order to persuade a group of people to converse about their museum visit. The system used cameras and microphones to capture the dynamics of group conversations and knowledge about the visit of each group member (the exhibits seen and the amount of information provided at every exhibit). While participating in the conversation, visitors’ attention was attracted by images or text displayed on the table.

The visualization was inspired by a Zen pond with red fish, displayed at the center of the table. The stimuli were pictures of exhibits and notes with short text, which appeared floating on the water during the conversation. An increase of fish activity swimming close to a stimulus or drops falling nearby were further means to attract the participants’ attention. The stimulus generation – adapted to the group rather than to the individuals – was based on the system’s understanding of the group behavior. Two inputs were used: participants’ attention estimated by a face detector processing the video from cameras, and their activity computed by a voice detector using the microphones. Several rules were defined for selecting appropriate content to present, according to the evidence of interest in a topic derived from the museum visit logs. For example, one rule targeted a participant, who was less engaged in the conversation, by providing a visualization of the exhibit she was interested in.

The system was evaluated by a Wizard of Oz study [14] and a user study with a working prototype [15]. The participants of the first study recognized that the table followed their conversation and the efficacy of the stimuli in fostering discussion. They also highlighted the need for reducing the number of stimuli and avoiding flashing or pulsing visualizations. The prototype study involved six groups of four visitors and was set up as a within-subject experiment, with random or adaptive visualization stimuli modes. The system initially visualized the stimuli in one mode and after seven minutes switched to the other mode, without notifying the participants. The results suggested that the adaptive stimuli were preferable in fostering conversations. The participants spoke a bit more in the adaptive condition rather than in the random condition, 38% vs. 34% of time. Also, the probability of the conversation being sparked by a stimuli was higher in the adaptive condition, 26% vs. 14%.

4.2 Discussion of the Personalization Dimensions

What facilitates personalization? The model of personalization was based on the observation of the behavior of users as a group. Although in a simple way, the system tried to understand the group dynamics and act upon them, to deliver indirect persuasive
Adaptivity and Personalization in Persuasive Technologies

strategies. The group model was elicited in an implicit manner – the system observed the group behavior to plan (and potentially adjust) the strategies. Moreover, the group modeling was dynamic, since it was updated continuously over the course of visitor interactions with the museum.

How is personalization achieved? The level of personalization employed by the persuasion strategies was relatively low; two dimensions were used: the target and the content. The former was decided upon an analysis of the group dynamics, while the latter was taken from the model of interests built for each user. The persuasion strategies were intended to change the ends, as the system suggested a topic to discuss. The intended behavior impact was short-term, although there was an implicit long-term effect of better learning and appreciating the museum visit. The persuasion was presented indirectly, by the means of visual stimuli in the pond. Finally, the personalization was not disclosed, in order not to interfere with the group behavior, although the Wizard of Oz study suggested that the participants still realized this.

How is personalization evaluated? In this case study, the evaluation was objective, since the system was assessed in a controlled study, by observing speech activity and topic shifts. Yet, for this type of intervention, it might also be important to consider a subjective evaluation measuring also the effectiveness (in terms of the behavior change), acceptability, and perceived intrusiveness of persuasion. It was a short-term evaluation and the metrics used were identical to those targeted by the intervention: speech activity and topic shifts.

5 Discussion

In this section, we revisit the above three case studies with respect to the dimensions of personalized persuasion that were listed in the introduction. Table 1 synthesizes the case studies. We discuss the chosen options, and review alternatives, along with their advantages and shortcomings.

Table 1. Mapping of the case studies onto the dimensions of personalized persuasion.

<table>
<thead>
<tr>
<th>Dimension</th>
<th>Case Study 1</th>
<th>Case Study 2</th>
<th>Case Study 3</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Features</strong></td>
<td>Level times</td>
<td>User responses</td>
<td>Group behavior</td>
</tr>
<tr>
<td>Implicit-explicit</td>
<td>Implicit</td>
<td>Implicit</td>
<td>Implicit</td>
</tr>
<tr>
<td>Dynamicity</td>
<td>Dynamic</td>
<td>Dynamic, not time varying</td>
<td>Dynamic, time varying</td>
</tr>
<tr>
<td>How is personalization achieved?</td>
<td>Tailoring and personalization</td>
<td>Low, simple</td>
<td>Low, two dimensions: target and content</td>
</tr>
<tr>
<td><strong>Granularity</strong></td>
<td>Means</td>
<td>Means</td>
<td>Ends</td>
</tr>
<tr>
<td><strong>End-means</strong></td>
<td>Short term</td>
<td>Short term</td>
<td>Short term</td>
</tr>
<tr>
<td><strong>Behavior term</strong></td>
<td>Short term</td>
<td>Short term</td>
<td>Short term</td>
</tr>
</tbody>
</table>
5.1 What Facilitates Personalization?

A range of features and user characteristics can facilitate the personalization of persuasive interventions. In the three case studies, personal characteristics (gaming skills, persuasion profile) and group behaviors (speaking, visual attention) are used. While in the first case a domain-specific feature was used, in the other two the dimensions are general, as the persuasion profile may be used in other domains. Domain-independence may be a desirable property in choosing features for personalization, since standard personality profiles may potentially predict the effect of persuasive strategies. At the same time, domain-dependent features may offer a deeper insight and allow more effective personalization.

Next, we consider whether user features are elicited explicitly or implicitly. In the explicit case, psychology research offers questionnaires and scales for profiling users with respect to various traits. The drawback of these is that in real-life systems, it may be unacceptable for users to spend time on the questionnaires. Although in some domains this may be acceptable and expected, the implicit approach, where user features are extracted from observable user behavior, is more popular. In some cases, like the first case study, the observations reduce the burden of users and possibly avoid bias, while in other cases, like the second case study, it may be inappropriate to ask direct questions.

Yet, even if the questionnaires are feasible and acceptable, they often cannot be administered too frequently, which highlights the benefit of the implicit approach applied continuously. This is likely to increase the accuracy of the user models and facilitates the delivery of a more fine-grained personalization. This dimension refers to the dynamicity and freshness of the models, and all three case studies opt for the dynamic modeling of the users.

5.2 How is Personalization Achieved?

The level of details or the granularity of personalized persuasion depends on the quality of the models. Detailed models that receive ongoing updates allow to adapt the service to the target users (or a specific group, as in the third case study) and provide fine-grained personalization. Contrarily, outdated or imprecise models may only allow tailoring of the interventions to groups or stereotypes of users. While the former is advisable, the latter may be the only choice, when a user cannot be accurately modeled. In the presented case
studies, the granularity of the models is low, which facilitates fairly coarse-grained personalization, primarily tailored to groups or stereotypes.

The second dimension refers to the actual part of the intervention that gets personalized: whether it is the means or the ends. Out of the three case studies, the first and the second present examples of means-persuasion, since in the first the interventions aim at increasing the number of jumps and in the second the end goal of each message is identical. In the third case study, however, each intervention has the same means (stimuli), but aims at increasing participation of a different participant. The combination of both is possible and in long-term systems, e.g., for health management, it may be advisable to personalize both the ends (target action) and the means (way of achieving this action).

The behavior term dimension focuses on whether the intervention aims at a short-term change in attitude or behavior, or at a long-term change in behavioral patterns. In all the case studies, as well as in the majority of related works, the impact is only short-term. Having said that, in all three cases, there is an underlying long-term goal (obesity control, healthy lifestyle, improved museum experience), but this goal is neither explicitly modeled nor targeted. Systems that explicitly aim at long-term stable behavioral changes are quite rare and they likely represent the next big challenge for personalized persuasive technologies.

Another dimension refers to the directness of the persuasion: whether the desired behavior or attitude is presented as the open target of the interventions or the interventions are masked by another activity. In the first case study, the target action pertaining to physical activity is masked by playing, such that the persuasive intervention is indirect. On the contrary, in the second case, the target action is stated in the messages and this is direct intervention. The third case is an example of indirect interventions, as the content of the messages is meant to foster conversation. It is also indirect in another sense: the pond is not meant to be the focus of attention and it is less intrusive than the messages.

Finally, the disclosure of personalization is an important dimension: to what extent should the system disclose that the interventions are personalized? The three case studies offer examples of undisclosed intervention. The third case study, however, suggests that the users understood the persuasive intent of fostering conversation. To some extent, persuasion should not be disclosed, in particular in indirect interventions, not to spoil its effect. Yet, it is important to acknowledge this in the system’s terms-and-conditions, which may rise ethical concerns. Privacy concerns may also arise; how is the collected user data man-aged and whether this information can be accessed by untrusted parties?

5.3 How is Personalization Evaluated?

The third aspect is the evaluation of personalized persuasion. The first dimension here deals with the objectivity of the evaluation. The subjective approach usually involves users assessing the system through questionnaires and interviews measuring their perception of or responses to persuasion. Contrarily, the objective approach consists of
obtaining indirect measures through analyzing system logs and user actions. The three case studies mainly adopt objective evaluation that is more appropriate when the interventions are not disclosed. Yet, in the third case study, user interviews helped to understand what worked and what did not.

Objective evaluations usually exploit a variety of metrics computed from the system logs and measuring user behavior, which is the ultimate target of the persuasion. As clearly comes through the three case studies, the metrics in place strongly depend on the behavior targeted by the persuasive interventions, and may differ immensely across systems. Yet, if the measurements are not done in a controlled environment, it might be contentious whether the change should be attributed to the interventions or to other factors. In this respect, subjective evaluations may be more informative, even though relying on self-reporting and willingness of users to provide feedback.

Finally, the evaluation term is an important, although often neglected, dimension. The majority of works, including the three case studies, conduct short-term studies and maintain simple experimental protocols. Findings from such studies may shed some light on the efficacy of the persuasive technology, but they often cannot be generalized and projected onto real-life applications. It should be mentioned that although long-term studies are expensive in terms of user involvement, robustness of the technology, and data collection methods, they are paramount for the evolution of personalized persuasive technologies.

6 Conclusion

This paper aimed at highlighting the choices that designers of personalized persuasive technologies face, when designing persuasive systems. Through three case studies and an explicit discussion of the dimensions, we have highlighted various options, discussed their advantaged and trade-offs, and motivated the design choices. Of course, the design space of personalized persuasive technologies spans beyond the above list of dimensions, and even for these dimensions the current state of research only partially informs the design choices. We believe, however, that our current analysis contributes to the ongoing work on better understanding the challenges and possible solutions in designing effective personalized persuasive systems.

References