PEACOX – Persuasive Advisor for CO2-reducing cross-modal trip planning

Project Reference: 288466

FP7-ICT 2011: 6.6 Low carbon multi-modal mobility and freight transport

Project Duration: 1 Oct 2011 – 31 March 2015

D7.3 Persuasive Strategies Evaluation Report

AIT

Author(s)

Johann Schrammel (AIT)

Sebastian Prost (AIT)

Efthimios Bothos (ICCS)

William Brazil (TCD)

Brian Caulfield (TCD)

Date: 27/02/2015

Dissemination level: (PU, PP, RE, CO): PU

Project Co-Funded by the European Commission within the 7th Framework Program

Abstract

This deliverable summarizes the results of different evaluation activities and studies targeting the effect of different persuasive strategies explored as part of the PEACOX project. The report provides a short introduction for each studied strategy, then describes the evaluation method and summarizes the findings and conclusions for each strategy.

Table of Contents

1. Introduction

Different persuasive strategies have been developed, studied and applied within the PEACOX project. This deliverable describes the overall approach and underlying assumptions towards our research of persuasive strategies, introduces a conceptual method to study persuasion and summarizes the findings of the evaluation of the different strategies applied and explored within the PEACOX projects. We provide discussions for future directions of research with regard to the application and further development of these strategies in the context of mobility. Some of the strategies were studied and formally evaluated in dedicated experimental setups; others were studied together with the integrated PEACOX system as part of the two field trials in Vienna and Dublin. Additionally results also consider informal feedback and comments on persuasive strategies collected during different evaluation activities of the project provided by users and experts.

2. Conceptual Approach

Persuasive interfaces have been proposed and used to influence people's behaviour and attitudes within the past couple of years. However, research found that such systems typically only result in small effects and that approaches which increase the effectiveness of persuasive systems are needed. A promising possibility to increase the impact of persuasive systems that is followed within PEACOX is to personalize and tailor the system concept, design and interventions to the individual user, context and target behaviour. In order to better understand the different interlinking of factors the following framework characterising the situation was developed within the project.

The suggested framework operates on and between three layers (see Figure 1 for a graphical representation). First, on the target behaviour layer, the system implements a number of persuasive strategies. Next, on the user layer, it personalises its services. Finally, on the situation layer, it adjusts to a specific context. The selection of strategies, personal, and context variables needs to be based on the system's application domain.

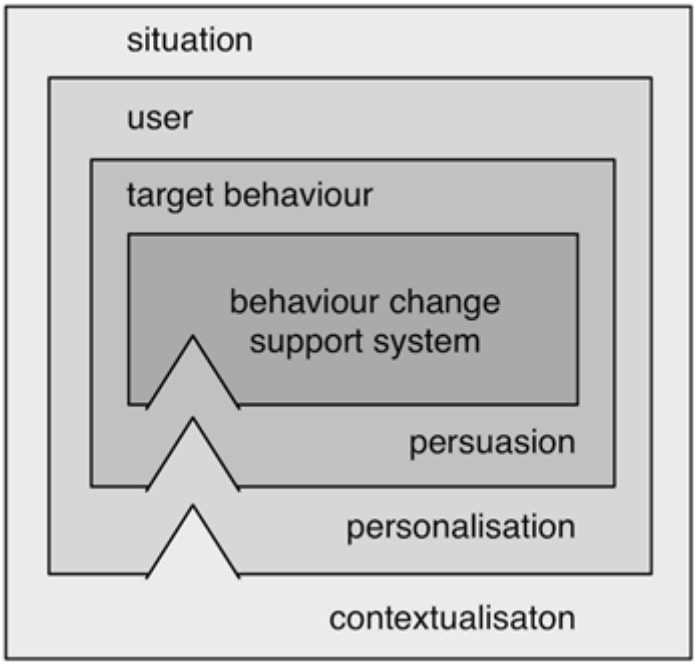

Figure 1: Persuasion Framework

On a target behaviour layer, different persuasive strategies and system designs have been suggested to support more sustainable behaviours. The target behaviour that is proposed in persuasive suggestions for behaviour change is tailored to an approximation of the individual users' range of acceptable travel alternatives which is derived from the users past behaviour and trip history.

On a situation layer, context-aware mobile computing has recently taken large leaps forward thanks to sensor-rich smart phones that allow systems to adapt to various situational variables.

On a user layer, personalisation has been identified as a key factor for travel mode choice and behaviour change. Even though personalization in persuasion has been suggested frequently, only few systems actually implemented personalization mechanisms, and limited empirical data on their effectiveness is available. Within the project we developed methods to estimate the persuadability and susceptibility to different persuasive strategies of users, thereby enabling us to select the most appropriate strategy for each type of users. We also started to analyse the influence and effectiveness of different persuasive agents in order to further personalize and individually target persuasive interventions.

As represented in Figure 1, the individual layers are not isolated but interlinked. Designers should therefore consider the following connections between layers:

- Personalisation should adjust to specific contexts.
- Persuasion should adapt to specific users and situations.
- The Behaviour Change Support System (BCSS) should respond to specific target behaviours, users, and situations.

For each layer, several behaviour change support factors (BCSFs) may be identified, depending on the specific application. It is those factors that interlink the different layers. They describe which factors on one layer influence which factors on another layer. For example, a certain situational factor (such as location or weather) can alter the system's personalisation and persuasion mechanisms.

In accordance with this model evaluation activities regarding persuasive strategies within the project addressed very different aspects, and a systematic and complete evaluation of the different factors was not possible. Therefore aspects and questions of specific interest and with promising potential for increasing the effectiveness of persuasive approaches were selected and further studied.

In the following chapters we therefore summarize the main findings of the different evaluation activities. Specifically we will report findings on the following topics:

- Measuring Persuadability
- Design of C02-Feedback
- Route Recommendations
- Persuasive Messages
- Challenges
- Temporal Dynamics
- Social and cultural factors limiting impact of persuasive strategies

3. Measuring Persuadability

An important development and evaluation activity focused on understanding the aspect of persuadability of persons towards different persuasive strategies. The goal of this aspect of work within PEACOX was 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 below). 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, an assessment of content validity of the items, the collection of data to conduct a pretest and the scale purification and refinement.

First we cover the scale development process, 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 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 (out of Torning & Oinas-Kukkonen 2009'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 fulfil 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 fulfilment 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: Fulfilment of the criteria (+1) or no fulfilment of the criteria (-1). In case of uncertainty, items did not have to be rated. 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 intended minimum of items per scale was defined as 6, to ensure an appropriate reliability a priori. 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).

In the second part of the scale development process, a item analysis (reliability by means of internal consistency and item-scale correlations) was performed, and based on the result the scale was refined. 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 analysed 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 more susceptible to these specific persuasive strategies. Therefore, these persuasive strategies have a greater effect for these participants than persuasive strategies where they have lower scores.

Table 1: Persuadability Questionnaire

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

Our analysis of scales for measuring persuadability has shown the successful initial development (creation, expert validation, item analysis by means of internal consistency in an online study) of five scales for measuring persuadability. Two scales did not show sufficient internal consistency and some items showed low item-scale correlations. The five 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: For example, Kaptein and Eckles have shown in 2012 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 challenge in improving these meta-judgemental measures until they can fruitfully complement or even compete with other source of data. With this work we want to stimulate discussions, ideas, approaches and more studies on this subject.

4. C02-Feedback

In this section we examine methods of communicating and presenting information to individuals about transport and travel related carbon emissions for online journey planners and smartphone applications with the goal to optimize the persuasive impact of this feedback. We first report the findings of a dedicated questionnaire study addressing design issues, and then discuss the feedback from participants in the second PEACOX field trials, where the concept was applied in praxis.

4.1 Questionnaire study

This study examined four methods of framing transport related emissions and examined the effect of these frames on ease of understanding and the potential to alter the respondent's mode of transport. Numerous on-line carbon calculators provide users with information about the carbon emissions that result from the selection of one mode of transport over another. Each of the methods used reflected an approach currently used by on-line carbon calculators.

To examine the research questions a survey was undertaken to assess user requirements for a persuasive travel advisor with the aim of reducing travel related CO2 emissions. This survey was conducted in Ireland in the form of an on-line questionnaire distributed via a number of sources including the electronic notice boards of Irish semi-state organizations. 457 responses were received in total with a 10 completion rate of 77.6%. 11. Due to the approach taken during the distribution of the survey, the sample is not representative of the Irish population as a whole.

Survey respondents were presented with four methods of understanding carbon emission arising from their trips. Each method presented the respondent with information on the attributes of the three modes available bus, driving and heavy rail. As the purpose of this study was to examine how carbon emissions information could be integrated into a smartphone application interface, information on travel times and trip costs associated with each mode were also presented as these attributes are likely to be included on any transport related application. These methods were based upon methods already being employed by journey planning applications and carbon calculators.

Method 1: the “Basic Numerical Method”, presented respondents with simple numerical information regarding the emissions that would be produced by each mode. Emissions information was presented in terms of mass in kilograms of CO2 produced by each mode with no additional information available to the user.

Method 2: also known as the “Light Bulb Method”, contained the same information as provided in Method 1 as well as additional information designed to help respondents put their emissions into context. Respondents were told how long a 60 watt incandescent light bulb would need to be left turned on to produce the equivalent amount of emissions of CO2 as their trip.

Method 3: known as the “Carbon Budget Method” presented respondents with the same basic information as provided in Method 1 as well as additional information regarding a daily carbon budget. The principle advantage of this method that it provides the user with a frame of reference that may not otherwise have been present. To avoid a bias in terms of framing effects respondents were told what percentage of their daily carbon budget each mode would consume per trip and also what percentage would remain.

Method 4: known as the “Traffic Light Method” was constructed in such a manner that while it contained the same information as the previous three methods with regard to travel time and trip cost, it omitted information on carbon emissions. This was intended to test whether respondents had a preference for visual rather than quantitative information on carbon emissions. Instead of numerical information, Method 4 provided respondents with a traffic light colour coding system where the highest emitting mode was assigned a red light, the medium mode a yellow light and the lowest emitting mode a green light.

Figure 2: Different methods of showing CO2 Feedback

After viewing the four methods of carbon presentation respondents were asked to indicate which method they had found the "easiest" and "hardest" to understand and which method was "most likely" and least likely" to entice them to move to a lower emitting mode. This question format forced respondents to make a choice between methods, while also addressing some of the issues of response similarity that may occurs with Likert scales.

Basic Numerical Information provided by Method 1 was deemed to be both the easiest understood and the most influential method with scores of 37.7% and 32.7% 48 respectively. Responses for the Light Bulb Method are very similar for both understanding and influence with scores of 28.2% and 28.5% respectively. The largest variance between understanding and influence was for the Carbon Budget Method where 17.7% of respondents stated that it was the easiest method to understand but 24.5% chose it as the most influential method.

The Traffic Light Method has been selected as both the method that is hardest to understand and least influential by largest section of respondents with scores of 40.3% and 47% respectively. The Carbon Budget Method was chosen as the hardest to understand by

29.9% of respondents and chosen as the least influential by 21.4%. This suggests that the Carbon Budget Method may be perceived as more influential than understandable. An implication of this may be that if users can be educated to operate a carbon budget system it may offer a more effective method of communicating and therefore controlling carbon emissions.

4.2 CO2 feedback in the field trials

Informed by the results reported above the PEACOX app also used numeric feedback for CO2 information, and feedback from users supported the importance of this approach. The CO2 values that PEACOX displays along the different route options turned out to be one of the most important persuasive elements in the system. It helped users to become at least aware of the impact of the different modes of transport. However, while all participants noticed the display of emission data, not all were influence by it. The detailed analysis of the trial results are provided in *D7.5 Field Trials II Report*, but in summary one can conclude that the findings in the trial very much reflect the findings from the questionnaire study.

5. Route Recommendations

Using targeted route recommendations was also evaluated as part of the first PEACOX field trial. The evaluation approach was user-centric and was based on the evaluation framework for recommender systems developed in Pu et al., (2011). More specifically, we gathered feedback on the following aspects, which provide the user perceived quality of the recommendations:

• Accuracy of the suggestions.

• Satisfaction with the suggestions.

• Diversity of the suggestions with respect to eco-friendliness: users were asked to state if they received eco-friendly routes.

• Novelty: users were asked to state whether they received eco-friendly suggestions they were not aware of.

Furthermore we gathered feedback on the following aspects related to the application usage:

• How often was the application used to plan routes and find specific routes.

• How often users followed one of the suggestions provided by the application.

User comments after two weeks of application usage revealed three issues. First, users that actively walked towards their destinations complained that the system displayed only short walking distance routes. This was due to the walk route filtering heuristic we were using which omitted routes with a walking time greater that 40 minutes.

Furthermore users noticed that certain park and ride routes did not make sense as the routing engine suggestion was to drive close to a location and then take public transportation for one stop. Last, for destinations close to the origin place, the routing engine was generating routes that were not usable, involving large circles with car or public transportation usage. All the aforementioned issues were handled before releasing a new version of our application. To this direction, we updated the filtering function to allow for longer walking trips, and we omitted park and ride routes if the usage of car and public transportation were disproportionally allocated. In order to identify strange routes for start and destination points in a short distance we implemented a mechanism which checked the shortest route and omitted the remaining routes that were too long compared to the

shortest. The questionnaire at the end of the trial revealed that users were engaged and were using often the application to plan their routes or search for specific routes (mean: 3.6; std: 1.2). Furthermore they stated that they followed one of the routes suggested by our application at least once (mean: 2.1; std: 0.8) and were satisfied with the suggestions (mean: 3.2; std: 1.1). When we looked into the user answers per transportation mode habits we identified that the most satisfied users by the route suggestions were those who stated that they normally walk (mean: 4;) followed by users who stated that normally use public transportation (mean: 3.7). The least pleased users were those who normally take their car (mean: 2.8) and cyclists (mean: 1.6). These results were in accordance with the question of how often users followed one of the suggested trips, where walk and public transportation users answered that they followed a trip more often. In terms of the application usage, we did not observe notable differences between these user groups.

Accuracy was perceived as good (mean: 3.1; std: 1.4) whereas users stated in the qualitative interview that the improvements we implemented positively impacted their feedback. When we asked users if they found eco-friendly routes within the suggestions, we received positive feedback (mean: 4.0; std: 0,8). Qualitative interviews revealed an increased environmental awareness, especially in cases of emission intensive behaviours. Car drivers reported a bad conscience when looking at the CO2 values for car trips, compared to public transportation, which was higher. Moreover, two users reported cases of switching from a bus to a tram because of lower emissions, and following a more eco-friendly car route compared to their usual route.

The results of our trial show that our approach was accepted by the users. Our implementation of default options did not annoy the users whereas they were pleased to receive information for trips with various transportation options to reach their destination. Users were also pleased to see the list of alternatives structured as they could make comparisons easily. Nevertheless, did not identify is a significant effect on users' behaviour. This is mostly due to the short time frame of our trial as well as our setting. Regarding the trial duration, a one month period seems to be inadequate to show any behavioural changes on transportation patterns.

Furthermore, our intention was to test the usefulness of our approach and not focus on the detection of behavioural changes. That is why we involved users with varying profiles, including individuals who favour the usage of car or public transportation or bicycle. The aim

per group of individuals varies. For persons who usually drive a car we want to apply behavioural changes. For persons who normally use public transportation and bicycle, we want to offer a useful tool which will help them to sustain these behaviours, and assist them in finding alternatives in times when they consider car usage. In order to observe the envisaged effects a longitudial study is required which will monitor users' behaviour for a number of months or years.

One of the findings suggests that our approach was not useful for cyclists. According to the comments we received cyclists' have specific needs from route planning applications and want to have functionalities in order to be able to identify scenic routes, safe routes, and routes with bicycle lanes. This kind of functionalities depend on the underlying infrastructure and specifically on the capabilities of the routing engine to find trips with such characteristics. In our case the routing engine lacked such functionalities and as a result we could not satisfy their needs.

Another lesson we learned is that relying on a routing engine for identifying recommendations of alternative trips to reach a destination is not sufficient. A higher level logic is required and needs to be implemented in order to prune trips that are redundant or do not make sense. For example, in our case, we had to implement filters to omit similar trips with different starting time, park and ride trips that involved usage of a car close to the destination and public transportation and trips with long detours although the destination was a few hundred meters away.

6. Persuasive Messages

The aim of the persuasive messages was to nudge users to follow routes with low emissions. Various persuasive strategies were employed whereas user profile attributes and contextual information determined the type of the message and the place of display (i.e. the specific route option) (see D5.5 for a description of the implementation).

In total 1558 messages were presented to the users in the second trial. Figure 3: Percentage of shown messages per modality in Vienna and Dublin.presents the percentage of messages shown in Vienna and Dublin per modality. The type of the message and the route in which it was displayed was affected by the user profile and the context of the request, thus the differences in the percentages. The (negative) message associated with the car routes was displayed the less times whereas messages associated with the public transportation routes were displayed the more times in both cities.

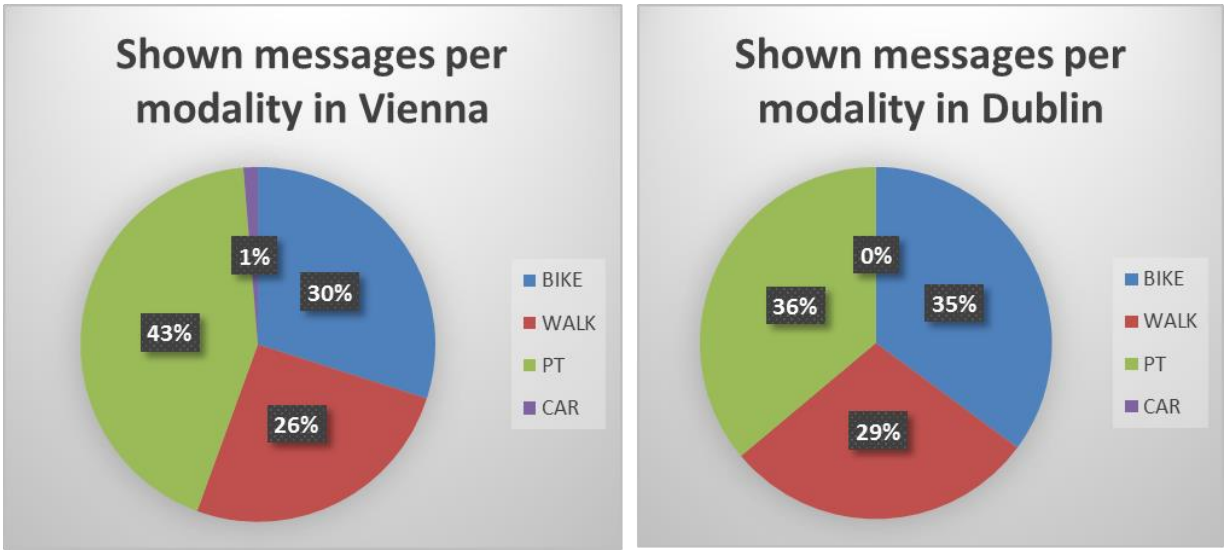

Figure 3: Percentage of shown messages per modality in Vienna and Dublin.

Figure 4: Messages and user selections. provides an overview of the user decisions in Dublin and Vienna. The first part of the text provided in each row of the table in the figure determines the main modality of the route where the message was displayed, whereas the second part the main modality of the selected route, e.g. the text “messagePTSelectedBike” means that the message was displayed to a route with main modality public transportation, but the user selected a route with a main modality of Bike.

Vienna			Dublin	
messageCarSelectedCar:	0		messageCarSelectedCar:	0
messageCarSelectedPT:	2		messageCarSelectedPT:	0
messageCarSelectedBike:	0		messageCarSelectedBike:	0
messageCarSelectedWalk:	0		messageCarSelectedWalk:	0
messagePTSelectedCar:	25		messagePTSelectedCar:	29
messagePTSelectedPT:	134		messagePTSelectedPT:	107
messagePTSelectedBike:	30		messagePTSelectedBike:	48
messagePTSelectedWalk:	7		messagePTSelectedWalk:	24
messageBikeSelectedCar:	13		messageBikeSelectedCar:	22
messageBikeSelectedPT:	82		messageBikeSelectedPT:	77
messageBikeSelectedBike:	37		messageBikeSelectedBike:	54
messageBikeSelectedWalk:	26		messageBikeSelectedWalk:	54
messageWalkSelectedCar:	6		messageWalkSelectedCar:	24
messageWalkSelectedPT:	61		messageWalkSelectedPT:	38
messageWalkSelectedBike:	24		messageWalkSelectedBike:	53
messageWalkSelectedWalk:	28		messageWalkSelectedWalk:	54

Figure 4: Messages and user selections.

After the trials also qualitative feedback was collected from the users (see D7.5 for details). Of the users that did consciously notice the messages, about 15% felt the messages do not always match with their current situation. Another 15% of users did find the recommendations realistic and useful, but nevertheless did not find them influencing their travel decisions. The remaining 20% of users considered the feature meaningful and also changed their decision at least a couple of times. In general users found the feature entertaining and made them smile. They appreciated the direct approach, as the user is addressed personally by the app.

7. Challenges

Another approach studied within the PEACOX project was the usage of challenges for persuasive purposes. Challenges provide users with a set of behaviour-goals, were they can voluntarily decide whether to accept or decline the challenge. This approach is based on the background of goal-setting theory (Locke &Latham, 1990). Typically such implementations are combined with social comparisons, which allow users to participate and engage in pre-defined challenges posed by some mediating instance (typically the provider of the BCSS) and to compare their success (and compete) with others. Whereas such approaches are common, surprisingly few empirical data on their actual performance and empirically funded guidelines for the design of such challenges exist. Questions like how to best frame these challenges, whether to better organize them in an individual or collaborative manner, how to tailor them towards specific user groups still remain largely unanswered from an empirical perspective.

We addressed some of these aspects in our work, and report the findings of an empirical study which researches the influence of different types of challenges and the user's characteristics on the effectiveness, perception and acceptance of challenges.

Challenges are closely related to the concept of goal-setting, and can be understood as a special case thereof. Goal-setting theory was originally developed within an organizational context and showed that specific and hard goals lead to better performance results than easy and unspecific goals. Especially the question of how to define the goal is of major importance. Building on goal-setting theory Consolvo et al. (2009) explored the preference of users on different goal sources (self-set, assigned, participatory, guided or group-set) and goal timeframes (fixed weekly scheme versus rolling time-window). Unfortunately no actual behavior data is available in this study, but an analysis of self-reported preferences indicates that self-set goals are preferred, but also that interesting design opportunities for guided and group-set approaches exist. Besides source and timeframe also the role of different defaults in the goal setting process has been researched recently (Loock et al. 2013). In the context of energy saving in the household the study found that default goals can lead to significant savings, and that it is important to choose the right defaults as both too low and too high goals have detrimental effects on the behaviour.

If the goal-setting is not done directly by the user, but the system presents the user with goals to achieve this is typically labelled challenges. Similar to classical goal-setting for challenges it seems to be important to achieve the correct level of difficulty as well as a close match to the user's intrinsic goals (Jylhä et al. 2013).

Using teams rather than individuals has also been suggested to improve persuasiveness (Staats et al. 2004). Consequently, team-based challenges are used in several systems, however only very limited data on their performance or guidelines for their design exist.

In detail we wanted to answer the following main research questions with our work:

a) Are challenges a successful means to sustain interest and motivation of users for engaging with behaviour change support systems?

b) Is there a main difference in whether challenges are framed on an individual or a collaborative level?

c) How do different types of individuals react to different types of challenges, and what are the design implications of such differences?

These questions were studied as part of the second field trials, where collaborative and individual challenges were organized, and data regarding usage, success and subjective experience was collected. A detailed description of the overall trial procedure, participants, methods and results can be found in D7.5.2 Field Trials II Report. In this deliverable we focus exclusively on the evaluation of the challenge strategy for persuasion.

As mentioned above there were two different types of challenges used in the study. In the individual challenges participants had to achieve a defined goal on their own. In collaborative challenges participants could join a group, which had to achieve the defined goal together, with each participant contributing to the overall goal. Altogether there were three pairs of challenges; these pairs differed from each other only in the individual or collaborative aspect. All six challenges were presented to the participants.

Challenges were designed so they can be achieved by all mobility types. The following three challenges either framed in an individual or collaborative way have been used in the trial:

Challenge #1: Identify 2 (for individual) respectively 10 (for collaborative) specific possibilities to save CO2 in personal transport. Please post your findings to the Facebook group.

Challenge #2: Try to lower your (individual or collaborative) CO2 emissions by 10 percent compared to last week or try to reach a value below 20g/km.

Challenge #3: Try to increase your (individual or collaborative) kilometres for cycling and walking by 10% or try to walk or cycle 50% or more of your kilometres.

Participants were instructed that taking part in the challenges is voluntary. To announce the challenges to the participants, we set up Facebook groups and events and asked the participants to join them. To counterbalance the sequence of individual and collaborative challenges we created four Facebook groups (two for each study site), one starting with an individual challenge and one starting with a collaborative challenge. The challenges were posted as Facebook events in the groups. Additionally participants got a notification through the app with a link to the Facebook event.

During the field trial, every week (except in week one and five) a challenge was proposed to the participants. Each challenge lasted for five to seven days. The travel behaviour of the participating users was analysed in the middle and end of the challenge period. In the middle of the challenge period, we presented an intermediate result to the participants by posting in the respective Facebook event site. For each challenge we measured how many participants took part in them and how many succeeded.

All participants who wanted to take part in a challenge had to accept the event invitation. As a reward for succeeding in a challenge participants could earn points, which defined the reputation level of each participant. By winning challenges the participants could rise from “wannabe” to “eco guru”. In order to support motivation of participating in the challenges information regarding their status was made visible on the Facebook group page.

Overall the challenges were perceived well by the participants, and a substantial amount of users participated in the challenges, although participation was voluntary. However, there seems to be a subgroup of users which did not like the challenges. Of our 37 participants 14 did not take part in a single challenge. Possible reasons for this behaviour are analysed in the section discussing factors influencing willingness to participate on the next page. Table 1 below provides an overview on the number of users that participated in the challenges over the duration of the field trial split according to individual versus collaborative challenge, and also provides information on how many of the accepted challenge were actually won.

Table 1. Participation and achievements in challenges

	Individual				Collaborative			
	Participated	**%**	**Won**	**%**	**Participated**	**%**	**Won**	**%**
Challenge #1	13	35.1	11	84.6	10	27.0	7	70.0
Challenge #2	15	40.5	8	53.4	12	32.4	9	75.0
Challenge #3	10	27.0	5	50.0	16	43.2	4	25.0
Total	38	34.3	24	63.2	38	34.3	20	52.6

Interest in participation in the challenges fluctuates somewhat (as has to be expected considering the limited number of users), but generally seems to continue over time and does not decline as might be expected. This is especially remarkable, as the app usage however slightly drops. This seems to indicate that challenges have the potential to better stimulate the long-term interest of users.

This interpretation is also supported by qualitative feedback collected in the inter-views. Several users stated that the announcement of challenges worked as a reminder for them to keep using the app. As one user puts it: *"It is motivating when you get 5 or 15 points as a reward. This is a motivational boost to keep using the app"*.

Moreover, the challenges helped some users to spur goals for greener mobility: *"I'm trying to fit it in, to exercise, to see how much more exercise I can do compared to using the car. The challenges were enabling me to use the bike. It's good to do exercise and take part in a challenge"*.

Overall most participants liked the idea of the challenges, as they were seen as a way to raise awareness and rethink existing behaviour patterns. The challenges were seen as an aid to make new experiences by pushing you to try out alternative modes of transport or routes: *"[I participated] because it was fun […], because one is motivated to do things different, stop doing things by the book. […] You are encouraged to try something new. [Normally,] you just do these things out of habit"*.

Also, several users liked that the challenges posed also personal goals: *"It was something to aim for. It's always nice to have targets, […] because you can't achieve anything without goals. So I tried to set myself little objectives, […] to be proud of achievements"*.

Another question of our study was whether individual versus collaborative framing of challenges influences the willingness to participate in them. Table 1 above shows the

average number of challenges the participants took part in. As is already obvious from the identical means paired samples t-test does not show any difference at all (t_{36}=0, p=1.00). Participation in collaborative challenges was highly correlated with participation in the individual challenges (r=0.833).

We also explicitly asked which type of challenges the users preferred. 14 stated that they did not look at the challenges closely and therefore cannot answer the question. About half of the remaining users did not have a preference (8) or didn't notice the difference (4). 8 participants clearly preferred individual challenges, and only 3 users favoured collective ones.

A very similar trend is present in the qualitative data. While during the interviews users were not explicitly asked about the two different types of challenges, none made an explicit statement differentiating between the two. All statements mentioned above refer to the challenges in general, not a particular type. When referring to challenges, many participants did, however, implicitly express a preference to individual challenges. Users were complaining that in group challenges other participants did not show enough engagement to complete the challenge: *"[Regarding] the challenges with ideas […]: Most of the suggestions in the group work are mine. I would have wished more collaboration from the others."* Also, the overall participation in the challenges was perceived as low by some: *"I'm surprised so few participate. For me, if I say yes to something I say yes to the whole of it."* This behaviour can be explained by the fact that the participants mostly did not know each other before the trial and therefore social pressure to work together as a group was low. As has been suggested before [24], collaborative and competitive mechanisms work better if participants know each other: *"There is another guy that I know who is participating in the study. So, I'm only just checking to make sure I'm ahead of him. […] There's lot of people […] I wouldn't know, so I'm not that interested in them"*.

We also analysed the success rates of the two types of challenges. On average the individual challenges were successful in 60% of cases, whereas the collective ones only were successful in 49.21%. This difference however is not statistically significant (paired-samples t-test with 18 users that participated in individual and collective Challenges, t_{17}=0.772, p=0.451).

Due to the fact that most users did not distinguish between the two types of challenges, and no clear differentiation is visible in the data, in the following analysis steps we do not distinguish between the two types of challenges anymore.

In order to further explore which factors actually underlie the users willingness to participate in challenges we analysed the data using a multiple linear regression approach with the number of challenges participated as independent variable. As dependent variables we used basic demographics (sex, age), ICT-competence, Environmental concern (measured using the scale provided in [18]) and a score of importance of the dimension of social comparison (based on [17]). This resulted in an overall marginally significant model ($F_{5,31}=2.326$, $p=0.066$, adjusted $r2=0.156$). As shown in Table 2 below only ICT-competence had a significant influence on the willingness to participate in challenges. The more competent users were the more likely they are to participate in challenges.

Table 2. Summary of multiple linear regression model

	B	Std. Error	t	p
Constant	1.539	3.296	0.467	0.644
Age	0.015	0.027	0.563	0.578
Sex	-0.686	0.740	-0.927	0.361
ICT Competence	-1.790	0.630	-2.839	**0.008****
Environmental Concern	0.565	0.527	1.071	0.292
Social Comparison	0.548	0.402	1.364	0.182

Some users remarked that the reward scheme through points and achievement levels was motivation for them: *"There were no real rewards waiting, but you could get virtual points, a form of reward. […] I thought that was fun"*. Also, not gaining points can be motivational factor: *„I didn't take part in the first [challenge]. I didn't get any points in the second group challenge. Then I thought, that can't be it, can it? And so I tried in the third and fourth if it is possible to achieve a little something using [different] modes of transport".*

Others did not share this excitement about virtual rewards. Just collecting points was not motivation enough. They would have wanted "real" value in order to change their transport behavior: *"Which grown-up person changes his behavior because of four points?"*. A number of users also explicitly expressed their disinterest in the game-like aspects: *"I'm not the player type. This has never interested me"*. Also, the competitive character of challenges was sometimes rejected: *"I'm not the competitive type. When I do something then I do it. Not because I want to trump someone"*.

Reasons for not participating. Participation in challenges does require users to commit additional effort. A number of users stated that they did not find the time to participate in challenges. Some of them stated they rarely use Facebook and therefore missed the challenge, despite the app notification that was sent out. For most users, the reason why they wouldn't find the time for the challenges, was that they were simply too busy in their lives. Others, however, blamed their own laziness: *"To be honest, I was too lazy. It was not immediately obvious how this works. There was too much to look at for joining. […] I thought I can't change my modes of transport anyway, because I will not cycle to work"*.

These findings suggest that challenges have the possibility to sustain the interest in using persuasive systems for a longer time compared to approaches that rely merely on feedback strategies. In order to maximize this effect known design principles especially the correct match of challenge difficulty and users' needs should be applied. Also, timing is important, and measurements of a users' activity level should be considered to decide when to prompt challenges to users. Alternatively, the possibility to postpone challenges in a simply manner could be a good design solution to allow the users to better fit challenges into their busy lives.

Our data suggest that individual challenges are the more appropriate means for organizing challenges, if there is no intrinsic collaborative aspect present. This is mainly linked to the users' possibility to feel in charge of the outcome (Self-efficacy). However, this finding does not mean that one should approach challenges in an isolated manner. Sharing of results of individual challenges was perceived as supportive, and has been shown to can have positive effects when designed well [8].

Participation in challenges was correlated with ICT competence. This is surprising, as a selection criterion for participating in the study was regular use of an up-to-date Android phone, which excluded technology-avoiding users. Even though the system was easy to use according to the feedback from our participants, users confident in dealing with new technologies seem to be more willing to engage in digital challenges. Consequently, it is even more important to design the challenges as user friendly as possible. Also providing different access means for users might be a good design solution, as this allows users' to rely on known technologies.

Another unexpected find was that environmental concern was not related to the number of challenges the user participated in. Concerned users might focus their energy on changing

their lifestyle, and playful means such as challenges are not needed to support this as they already are willing to do so. However, more data and research is needed to confirm (or disprove) this conjecture.

Some users explicitly rejected challenges as they dislike games and competitions. Therefore designers need to carefully consider their use. Gamification elements should be voluntary, so that users can simply ignore them. Regarding the persuasiveness of a system, elements like social comparison, competition and virtual rewards should not be the only strategy to be included, as some users will not respond to them.

In conclusion we can say that challenges seem to present an important possibility to sustain interest in interacting with persuasive technology, even though this is only true for a subset of users. Further research is needed regarding the question of how challenges can be made more attractive to these users, or to identify other approaches that better address the needs of them.

Our results also suggest that individual challenges seem to be preferable over collaborative ones, especially in application contexts were the users are not personally known to each other, and where there is no intrinsic collaborative aspect of the task.

8. Role of temporal dynamics – How to calculate feedback over time

Feedback on various kinds of behaviour and its effects on the environment has been used to influence the users behaviour towards more environmentally friendly behaviour patterns in numerous systems. Feedback has been shown to be able to produce effects (even though they are typically small i.e. in the area of 5-10 per cent) in different application areas such as e.g. domestic energy consumption or CO2-production related to personal travel. Related research has shown that the way in which this feedback is expressed and the detailed design solutions are important factors contributing to the effectiveness of behaviour feedback. Comparison to historic self-behaviour has been identified as important factor for the design of eco feedback. Also, expressing the feedback in simple semantic terms rather than in abstract numbers and units has been shown to be easier to understand and more adequate for the users.

A number of feedback approaches that combine these two mechanisms have been suggested and implemented. A very common approach is to calculate how good a user is currently doing combined to his past behaviour, and to summarize this in form of a real-world metaphor, typical examples being a plant growing, an iceberg melting, etc. A key question for the design of such a system is the question of how to calculate whether a user is doing good or not, and how strong the penalty or reward should be in order to maximize the persuasive aspect of the design. One must not forget that the design goal here is not correct or fair calculation, but the maximization of behaviour change.

Another important aspect of these algorithms is the time frame they do consider relevant for comparison: is the whole past behaviour relevant, is the current behaviour compared to a specific reference period, or is a sliding time window applied.

Our main research question was how does or proposed algorithms evolve over time, and what its expected potential for persuasion is.

In order to evaluate the different algorithms we defined a set of criteria and desirable characteristics for self-comparison algorithms based on theoretical analysis and state-of-the-art:

- Does not become stuck on extremes
- Positive actions are immediately followed by perceptible rewards
- Individual bad behaviour should be able to recover (forgiveness)
- Keep providing meaningful feedback and possibilities to act in long term use

We then prepared data from the gps-tracks of the first field trials in Vienna (trip logs of 10 users for 8 weeks), and applied the tree visualization paradigm as the main visualization form. The resulting visualizations and especially their development over time then are analysed with regard to important key figures (e.g. variability, etc.) and in comparison to the above defined target characteristics. The table below provides an overview of the key development characteristics during the first trial:

Table 2: Development of persuasive tree over time

User Id	Total Distance (traced, in Km)	Improved behavior detections (# of times)	Worsened behavior detections (# of times)	Final percentage of Tree growth	Stated Profile
1	149657,8056	4	29	0,268941	car
2	4784585,223	25	8	0,663739	pt
3	291157,426	17	19	0,480011	car
4	663174,3649	24	6	0,672607	pt
5	31537,87081	9	7	0,519989	car
6	795276,1701	22	16	0,559714	car
7	57493,79654	23	8	0,645656	pt
8	5154208,598	22	18	0,539915	walk
9	102769,9213	32	8	0,723122	bike
10	5897,291588	7	7	0,5	pt
11	651306,0005	12	27	0,354344	bike
12	31186,80285	17	4	0,627148	pt

The table shows that the tree behaved as expected in general, and that most of the desired characteristics were met.

In the second trials also extensive qualitative feedback on the behaviour of the tree over time and the perception of it by the users were collected (see *D7.5 Field Trials II Report* for details).

Most of the users also paid active attention to what the tree is showing, and felt positive about it. The tree was successful in creating an emotional response to its growth or shrinking. Many participants hoped that their efforts are reflected in the tree, and if, for particular reasons, this did not happen, they felt disappointed. Most users did continue to monitor the progress of the tree throughout the study period. Overall, the tree can therefore be characterised as an unobtrusive and engaging way of keep participants interested in the app. While its persuasive impact should not be overestimated, it serves as personal and emotional reflection of one's actions and can create feelings of connection and responsibility towards its "wellbeing".

9. Social and cultural factors limiting impact of persuasive strategies

During the field trial it also became evident that many participants did not change their mobility behaviour. A grounded theory approach was taken for analysis of the interview data to answer the question, why this was the case. In a first step, we performed open coding to derive *factors* from the data that influence travel behaviour. This process resulted in 20 factors as listed in Figure 1. We then sorted these factors roughly by how much they are under an individual's personal control. Additionally, we grouped factors and labelled them as *spheres of influence*, suggesting they should not be seen as distinct categories, but rather as markers on a continuum. We defined 4 spheres: *personal, social values, societal,* and *structural*. In a last step, we developed broad *approaches* how sustainable HCI can address these factors.

Figure 5: Spectrum of social and cultural factors

The personal sphere contains factors an individual can directly change and where it is comparably easy to support such change with technology. For example, one user stated that he does not want to stop driving "mainly because it is fun. This is a very important point" (I1). Besides fun, "pure convenience" (I4), the acquisition of a driving license and habits are influential factors.

The social sphere contains social values that guide individual behaviour. Obvious values include environmental concern and status. One user put this bluntly: "I mean, excuse me, as a businessman, you drive a car. You just don't take the subway" (I3). Privacy is another

value. One user for example, prefers the privacy of his car, because "when you get on a tram with a dog, people just look at you funny." (I2). We also consider flexibility and saving time as social values, as the need for being flexible and time effective is not naturally given but culturally shaped.

Many drivers reported the need of the car to travel between work and home, and more generally to travel longer distances: "I work in [the outskirts of Vienna], and since I've moved to the inner city I have been driving much more often than before" (I6). A second, very common answer was to transport children and/or items: "As you know, when the kids are little there are many things to take with you, […] diapers, […], toys, clothing, and so on" (I4). Again, cultural values shape where one finds it desirable to live and work, e.g. influenced by where social peers live. Also the desire to transport children by car, be it out of convenience or out of safety concerns, is not inherent to humans.

Another important sphere is the societal sphere. Someone's financial situation is not a personal choice, in many cases it is a fact defined by society. One user reported, "As a student I was only riding the motorbike, because of financial reasons. Now I can afford both" (I1). Likewise, changing social roles that are associated with older age or retirement can be a factor for changes in travel behaviour. One participant, 68 years old, was a car driver all life long but stopped abruptly once being retired: "I think this really changes with age. Today I think this is really stupid to drive a car through Vienna" (I3).

The final sphere (Structural Sphere) contains factors that are structural, and thus very difficult to change for an individual. They are concerned with the reliability, safety and availability of public transportation and cycle paths. One user thinks that "public transport is partly overcrowded, and partly it has the drawback to fail often" (I5). On the other hand the same interviewee does not cycle, "because I do not necessarily feel safe on the cycle paths" (I5). And sometimes, as one user explains, there is no alternative to the car: "One thing is clear, if you only have a bus twice a day, and you want to go somewhere, there is no other option" (I7). At the very end of the continuum, the weather can be a reason for not cycling: "Well, sometimes it's the weather's fault" (I7).

Implications for HCI. As the results show, there are a number of factors preventing sustainable behaviour. The point of grouping them into *spheres* is to show that a large number of these go well beyond the individual's influence. In fact, it can be argued that everything outside the personal sphere would require larger social or cultural changes to

overcome. We therefore encourage an activist stance on dealing with the complex issue of sustainability, but do not condemn "conventional" approaches focusing on the individual. Figure 1 shows that there is space (and, as we argue, need) for a variety of ways to address the identified factors. In particular, sustainable HCI should provide *entertainment*, *education*, *community support*, and *political activism*.

Entertainment

To entertain means to increase the fun factor of sustainable alternatives to car driving. Gamification approaches such as Greenify (Lee 2013) seem promising. However, entertainment alone cannot fully cover all aspects of unsustainable practices. One important addition is to add educational aspects.

Education

Educational products have value when users should be informed about consequences of personal actions, optimisation potentials, and possible alternatives. Furthermore, they are implicitly normative by defining what is desirable behaviour and thus try to provoke self-reflection on personal values. There already exist plenty of educational products that implement this, both in research and commercially available. A famous example is UbiGreen (Froehlich 2009), and the PEACOX prototype also falls into this group.

Community Support

In addition to education, values can be addressed by social support strategies, such as social comparison, which has been a popular strategy in the field of persuasive technology. Examples include various carbon emission or energy consumption meters that allow comparing your own behaviour with others. Additionally, community mechanisms that support direct communication between its members to share ideas and tips, for example in health promotion (Kaptein et al. 2010), allow users to take a more active role, which leads us to the last group of approaches.

Political Activism

Activism approaches have taken different shapes in HCI. Prominent examples are citizen science or citizen sensing approaches, where people use today's sensor-rich ubiquitous technology to collectively gather large amounts of data, e.g. on air quality (Paulos et al. 2008). Important here is to empower users to be active data generators rather than passive sensors. For example, technology could enable them to better reach out collectively to political representatives or the media to improve situations, e.g. missing public

infrastructure or unsafe cycle paths. Even the weather does not need to be a barrier. Experienced cyclists could serve as agents of change for all-weather cycling, including tips on appropriate bike equipment, clothing, and cycling style.

Conclusions

In this section we presented results of an analysis of qualitative interviews that explored why users did not increase use of sustainable transport alternatives despite being exposed to persuasive technology that promoted such behaviour. The results enrich the on-going debate in the HCI community on the role of activism in persuasive sustainability. The point is made that "conventional" approaches fail to recognise a number of factors that strongly influence an individual's behaviour and are therefore likely to have limited success, if any. Furthermore, it is argued that activist approaches that have been proposed before by others are a fruitful addition to overcome this limitation. HCI can empower users to become activists for their cause to create wider social or cultural change. In contrast to other, more radical positions, we do not dismiss conventional approaches but see *HCI for activism* in symbiosis with them to address the full continuum of influence factors pictured in this article.

Admittedly, the list of influence factors presented in this work does not claim completeness. The full analysis of interview data collected will most likely reveal additional factors to be included in the continuum, or even open up other lines of thought. Furthermore, the order and grouping of factors needs to be validated quantitatively with a larger sample size. Nevertheless, the preliminary results allow a first empirical validation of the importance of activism in HCI.

10. Conclusions

In this deliverable we reported the results of different evaluation activities and studies targeting the effect of different persuasive strategies explored as part of the PEACOX project. We briefly introduced the different strategies and provided a summary of the findings and conclusions for each strategy.

11. References

Kristian Torning and Harri Oinas-Kukkonen. 2009. Persuasive system design: state of the art and future directions. In *Proceedings of the 4th International Conference on Persuasive Technology* (Persuasive '09). ACM, New York, NY, USA

Kaptein, M., Eckles, D.: Heterogeneity in the Effects of Online Persuasion. Journal of Interactive Marketing. 26, (2012).

Pearl Pu, Li Chen, and Rong Hu. 2011. A user-centric evaluation framework for recommender systems. In Proceedings of the fifth ACM conference on Recommender systems (RecSys '11). ACM, New York, NY, USA

Locke, E.A., & Latham, G.P. (1990). A theory of goal setting and task performance. Englewood Cliffs, NJ: Prentice-Hall.

Consolvo, S-., Klasnja, P., McDonald, D.W., Landay. J.A.: Goal-setting considerations for persuasive technologies that encourage physical activity. Persuasive '09, ACM, (2009)

Loock, C., Staake, T., Thiesse, F.: Motivating energy-efficient behavior with green IS: an investigation of goal setting and the role of defaults. Mis Quarterly, 37(4), (2013)

Jylhä, A., Nurmi, P., Sirén, M., Hemminki, S., & Jacucci, G. Matkahupi: a persuasive mobile application for sustainable mobility. UbiComp'13, (2013)

Staats, H., Harland, P., & Wilke, H. A. (2004). Effecting durable change a team approach to improve environmental behavior in the household. Environment and Behavior, 36(3).

Jon Froehlich, Tawanna Dillahunt, Predrag Klasnja, Jennifer Mankoff, Sunny Consolvo, Beverly Harrison, and James A. Landay. 2009. UbiGreen: investigating a mobile tool for tracking and supporting green transportation habits. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '09). ACM, New York, NY, USA,

Kaptein, M., Lacroix, J., and Saini, P. Individual differences in persuadability in the health promotion domain. Proc. Persuasive, ACM (2010)

Paulos, E., Foth, M ., Satchell, C., and Kim, Y. Ubiquitous Sustainablity: Citizen Science and Activism. UbiComp ’08, (2008)

Joey J. Lee, Eduard Matamoros, Rafael Kern, Jenna Marks, Christian de Luna, and William Jordan-Cooley. 2013. Greenify: fostering sustainable communities via gamification. In CHI '13 Extended Abstracts on Human Factors in Computing Systems (CHI EA '13). ACM, New York, NY, USA