D 5.3 FIRST VISUALIZATION AND FEEDBACK INTERFACES AND BEHAVIORAL GAME CONCEPT

Project title	**Collaborative Recommendations and Adaptive Control for Personalised Energy Saving**
Project acronym	**enCOMPASS**
Project call	**EE-07-2016-2017 Behavioural change toward energy efficiency through ICT**
Work Package	**WP5**
Lead Partner	**EIPCM**
Contributing Partner(s)	**POLIMI, KAL**
Security classification	**PU (Public)**
Contractual delivery date	**31/01/2018**
Actual delivery date	**31/01/2018**
Version	**1.0**
Reviewers	**PMI, PDX**

History of changes

Version	Date	Comments	Main Authors
DDP v1	07/12/2017	Draft DDP	K. Koroleva (EIPCM)
DDP v2	12/12/2017	Revision of DDP	M. Melenhorst (EIPCM)
DDP v3	15/12/2017	Final version of DDP	M. Melenhorst (EIPCM) J. Novak (EIPCM)
V0.1	18/01/2018	Overview Visualization Input	K.Koroleva (EIPCM)
V 0.2	22/01/2018	Worked in the comments of Jasminko	K. Koroleva (EIPCM)
V 0.3	23/01/2018	First version of the awareness app and backend	Piero Fraternali (PMI)
V 0.4	23/01/2018	Impact visualization input	K. Koroleva (EIPCM)
V 0.5	24/01/2018	Completing awareness app and Funergy game	Piero Fraternali (PMI)
V 0.6	26/01/2018	Further continue impact visualization input	K. Koroleva (EIPCM)
V0.7	26/01/2018	Minor changes, input from Mark	K. Koroleva (EIPCM)
V0.8	26/01/2018	Version ready for internal review	K. Koroleva (EIPCM)
V0.81	26/01/2018	Internal check and corrections	J. Novak (EIPCM)
V 0.82	29/01/2018	Introduction, conclusion	M. Melenhorst (EIPCM)
V 0.9	29/01/2018	Worked in comments, version for QC	K.Koroleva (EIPCM)
V 1.0	30/1/2018	Quality checked version	P. Fraternali (PMI), L. Grillo (PDX)

Disclaimer

This document contains confidential information in the form of the enCOMPASS project findings, work and products and its use is strictly regulated by the enCOMPASS Consortium Agreement and by Contract no. 723059.

Neither the enCOMPASS Consortium nor any of its officers, employees or agents shall be responsible or liable in negligence or otherwise howsoever in respect of any inaccuracy or omission herein.

The contents of this document are the sole responsibility of the enCOMPASS consortium and can in no way be taken to reflect the views of the European Union.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 723059.

Table of contents

1 EXECUTIVE SUMMARY

This document contains the enCOMPASS deliverable 5.3 First Visualization and Feedback Interfaces and Behavioral Game Concept, which according to the DoA comprises:

"First design of the mappings from consumption data to persuasive visual metaphors. First version of gamified awareness apps and game concept; integration of an existing back-end supporting game/gamification. Test with users and collection of feedback."

The developed visualization designs show how energy consumption will be visualized in an intuitive, easily understandable way that relates abstract concepts to well-known metaphors from everyday life. In addition to visualizing the consumption itself, the developed visualization model also shows how the impact of achieved energy savings can be visualized in a way understandable for the users. The developed visualization model is also adaptive in that it considers that different types of users have different motivational preferences i.e. goals for saving energy. The results of the crowd-based user tests confirm the suitability of the developed visualizations and persuasive visual metaphors, and the collected user feedback will be used to further optimize the visualization design. The behavioral game concept has also been developed (hybrid digital-card game), alongside with the corresponding visual design. The first prototype of the enCOMPASS gamified awareness app has been implemented, consisting of a first version of the client-side app integrated with a gamification engine back-end. For this first functional prototype an alpha test has also been performed with a small group of users, confirming the overall concept and design of the app and providing feedback for further development.

The deliverable is organized as follows:

- **Section 2** Provides a brief introduction.
- **Section 3** Describes the design rationale behind and the process of the development of visual metaphors, as well as provides an overview of the metaphors themselves, both for the overview and the impact visualization. It keeps the rationale presented in Deliverable *D2.2 Final Requirements* and builds on the Mock-Ups developed for the overview and impact visualization.
- **Section 4** Presents the first version of the integrated gamified awareness application and concept, shows the screenshots of the first application prototype and their intended purpose, describes the back end functions of the application, as well as presents the first ideas on the logic behind the gamification concept and the Funergy game.
- **Section 5** Presents the results of feedback collected through three tests: understandability of the overview visualization, understandability and preference for a specific type of impact visualization, as well as the first usability test of the prototype of the application. It presents a vast array of quantitative and qualitative findings and shows that overall the visualizations of the enCompass application were well received by the users, they create a tangible perception of energy savings and would motivate them to save energy.
- **Section 6** Draws the conclusions.

The relationship of this deliverable with the other ones can be summarized as follows:

- *D5.1 Behavioral change models and determinants for energy consumption* and *D5.2 Incentives and Engagement Strategies* have provided an extensive literature review that has yielded possible incentive mechanisms and behavioral antecedents that can be targeted by means of the visualizations designed and evaluated in this deliverable.

- In *D2.2 Final Requirements* the first designs of the visualizations have been developed and tested with end-users in the requirements workshops at all pilot locations. The designs were informed by the first results from wp5.
- The designs of the visualizations will be further refined, and subsequently implemented as part of the first encompass release (*D6.3 Platform Initial Prototype* (M18)).

2 INTRODUCTION

Consumption feedback is an important part of the incentive model in enCOMPASS. The ambition of the project to go above and beyond traditional visualization approaches by e.g. introducing adaptive elements, and combining both metaphor-based and traditional approaches (e.g. bar chart), requires an iterative process to optimize the designs for such consumption visualizations.

As the first iteration, in the requirements phase mock-ups were developed that offered three different views on the user's consumption (an at-a-glance overview, three different impact visualizations, and a traditional detailed bar-chart visualization). Notwithstanding minor improvements on specific design elements, the requirements workshops with end-users overall confirmed the suitability of this approach that was developed to incentivize different types of users, with different motivations for saving energy.

This deliverable takes the end-user feedback collected in WP2 as a starting point for the refinement of the visualizations. The different views are further improved, and iteratively tested with end-users, focusing on both the concept, and on user acceptance criteria (e.g. understandability). First, we test among users on a crowdsourcing platform, then the first integrated version of the awareness application is evaluated with a convenience sample that closely resembles the pilot users, as to get an early assessment of the visualizations in relation to the other elements of the incentive model.

After alignment with the technical partners, the presented visualizations will be updated, implemented, and integrated in the awareness app that will be released in M18.

3 DESIGN OF VISUAL METAPHORS

In this section we first present the design rationale, distinguishing between different consumption visualizations available from the awareness app, followed by the design process, and the preliminary visualizations that resulted from multiple rounds of end-user feedback, which will be further refined and checked for technical feasibility before implementation in the first release of the enCOMPASS awareness app in M18 (end of April).

3.1 DESIGN PROCESS AND RATIONALE

The design process of the consumption visualizations fuses insights from literature on consumption feedback in behavioral change applications for energy saving, with lessons learnt from the requirements analysis that were reported in *D2.2 Final requirements*. Drawing on the literature on behavioral change antecedents and incentives (D5.1 and D5.2, with preliminary results reported in D2.1), and initial interviews with stakeholders (D2.1), mock-ups were developed with the initial visualization concept and the division between different consumption views, which were formally specified in D2.2 in a user story and a set of use cases.

Subsequently, these preliminary mock-ups were evaluated with end-users in the requirements workshops. The feedback from the users, was employed to elaborate and refine the visualization concept for each of the metaphor-based views. Subsequently, the refined visualizations were iteratively tested with end-users, recruited from Amazon Mechanical Turk's crowdsourcing platform. As a final step for this deliverable, an alpha test was conducted among users closely resembling the pilot users. In this test, the integrated awareness app was evaluated, in which the detailed consumption chart was actively working, and the metaphor-based visualizations were integrated as screenshots.

The remainder of this sub section provides the rationale behind the design choices that were made over the course of this design process, while in Section 2.2-2.4, the latest versions of the visualizations are depicted and described. The test results are reported in section 4, covering both the tests on Amazon Mechanical Turk, and the alpha test of the integrated awareness app.

As a detailed theoretical underpinning of the behavioral change incentive model has been provided in D5.1 and D5.2, this section briefly outlines the specific design choices that were made for the metaphor-based consumption views and the data-oriented consumption chart. While research in environmental psychology in general, as well as specific studies on consumption behavior for natural resources suggests that feedback has the potential of influencing underlying beliefs regarding energy consumption and attitudes towards energy saving (e.g. Steg et al., 2014; Tiefenbeck, 2016; Novak et al., 2018), the design of visualizations is not trivial for a range of different reasons.

First, users have different environmental goals and values (Lindenberg & Steg, 2007), as well as different needs with regard to energy consumption feedback (Gölz & Hahner, 2016), among others because users differ in terms of their behavioral change progress (e.g. Bamberg, 2013). In contrast to existing approaches that take a 'one size fits all' approach, in enCOMPASS the visualizations in enCOMPASS are personalized, based on the user's motivations for energy saving.

Second, the structural characteristics of energy consumption behavior as abstract, non-sensory, comprised of multiple behaviors, and of low personal relevance to most individuals (Karlin et al., 2015) impose challenges on the designers of energy consumption visualizations, among which representation of energy data in meaningful units, as well as the temporal grouping of data (e.g real-time, by day, week or month) (Schwartz et al., 2015).

As qualitative research has demonstrated that the abstract units of W and kWH, and the distinction between the two, are difficult to understand for household users (Karjalainen, 2011), a metaphor-based visualization approach can be used, to help users cope with the cognitive load of interpreting complex numerical, and abstract information. Examples of such metaphors include the use of traffic lights and gauges (Monigatti et al., 2010; Rist, 2014; Sundramoorthy et al., 2010), or eco-visualizations (Froehlich et al., 2009; Gustafsson, 2009; Rist, 2014) that map energy consumption to objects from nature's (e.g. trees, sealife). Particularly in the case of enCOMPASS where visualizations are used among users of schools (e.g. young children), public buildings (e.g. employees), and households with varying characteristics, easily understandable information is crucial.

Finally, consumption feedback should be considered as part of a system of assistance designed to enhance user engagement with energy consumption (Fréjus & Martini, 2016). To be appropriable, it must anticipate user needs evolving over time, it should be functionally rich, containing multiple feedback options, it should be complementary to other forms of assistance, and it should include an estimate of the impact of the energy consumption (ibid.).

Research on the design of consumption visualizations has yielded several design principles that have been used for the design of consumption visualizations in enCOMPASS:

- Allow users to **compare their consumption against an (historic) goal** that is constantly clear to the user, that is tailored to the specific household, and that contains concrete actions for how to achieve the goal. (Jacucci et al. 2011; Micheel, 2008; Geelen et al., 2013);
- Allow users to interact with feedback (Fischer 2008), by using **layers with different levels of detail** (Micheel et al., 2014; Jacucci et al. 2011);
- Provide **normative persuasive messages** alongside with consumption feedback (Fischer, 2008)
- **Differentiate between users with different goals** (Gölz & Hahnel, Micheel et al., 2014), and different levels of data-affinity (Micheel et al. 2014);
- Feedback should be designed in an **easy-to-use, visual appealing** way (Jacucci et al., 2011; Micheel et al., 2008);
- Use **visual metaphors** related to the user's consumption context, to facilitate comprehension of abstract consumption information (cf. Ludden, 2013; Rist, 2014; Gustaffson, 2009).

From these design principles and design challenges, three different consumption views were distilled that are accessible as pages in the enCOMPASS awareness application. The three views offer different layers of detail and different representations of the consumption feedback, as to allow users with different levels of data affinity (Micheel et al., 2014), and different levels of intrinsic or extrinsic motivations for energy saving, to choose and switch between views that best fit their needs at that moment in time. Metaphors are employed in which the difficult to understand unit of kWh is changed to reduce the cognitive load inflicted on the user, and to facilitate comprehension (e.g. Ludden, 2013).

As defined in the requirements (*D2.2 Final requirements*), the views comprise:

The battery overview

The battery was chosen as a metaphor of a tangible object, well-known to users, which resembles a target amount of energy that depletes over time. In terms of the temporal representation of consumption information, the battery displays the current month's energy consumption, while users can navigate to previous months to check their eventual consumption over the past months.

The amount of available energy is based on the baseline consumption from the same month last year. Users can set their goal relative to their baseline consumption, which is by default 20% less than last year (the KPI level), but can be changed to more or less ambitious levels. From both psychological research and research specifically in the area of feedback systems, it is known that difficult, specific, context-appropriate, and immediate rather than long-term goals, motivate to achieve more (Ling et al., 2005; Locke et al., 1981).

Consumption alerts are provided as color-coded normative messages, which display the user's progress in relation to the set goal. A green message appears when the user can still achieve his/her goal, including a motivational message to do so. Orange is used when the user is close to *not* meeting his goal (e.g. Be careful!), while red is used to state that the user cannot achieve his/her goal anymore (e.g 'You're using too much energy'). As consumption feedback should be actionable (e.g. Micheel et al. 2014), tapping on the 'Learn more' button will directly take the user to the tips page, where the user can learn how to save more energy. To allow users to track their consumption within the month, and to observe differences between the weeks, every week's consumption is displayed with a dashed line, which can be toggled on or off.

Graphical representations of the goal, and the baseline consumption, and the current consumption within the battery are supplemented with numerical information (e.g. the respective number of kWh's, and the difference with the baseline in %).

The impact view

The impact view is personalized based on people's main motivation for saving energy, drawing on Goal Framing Theory (Lindenberg & Steg, 1997), and the motivations for using energy feedback systems (e.g. Gölz & Hahnel, 2016). For each of these main motivations, a separate impact view has been designed: saving money, protecting the environment, and having fun while saving energy. While users have a focal motivation for saving energy, in line with Goal Framing Theory, all motivations are expected to be active at the same time, to some extent. To cope with these multiple active goals, users can switch back and forth between different impact views. However, the user interface is adapted based on the user's answers to questions about his/her motivations in the sign-up form (see also Section 3), whereby the most important goal is used to select the impact view that is displayed by default. By analysing user behavior data during the first release period, prior to Release 2, a more behavior data-driven approach can be implemented to select the view that is displayed by default.

Each impact view represents the last year's savings, starting from the current month. Thus, savings accumulate over time, which appeals both to the direction 1 desire of collection (Reiss, 2014), and affords a sense of achievement (cf. *Need Achievement Theory*, Atkinson & Litwin, 1960). The intended effect of the user inspecting the accumulation of his savings over time is fostered by allowing users to inspect the total savings over the past months, using similar buttons as in the consumption battery overview.

For each view, two metaphoric units are chosen that relate to the energy saving motivation. The bigger unit is filled over time with smaller saving units, until it is full, at which moment a new bigger unit is chosen.

The *'save money' view* displays the total amount of money saved on the electricity bill. For the 'save money' view, the piggybank was chosen as the bigger unit, while the smaller units with which it is filled are coins. As has become clear from both the requirements analysis (D2.2), and from literature, users have shown to appreciate visualizations of money spent or saved (e.g. Karjalainen, 2010), in spite of concerns about long-term impact on the user's motivation to save energy (Spagnolli et al., 2011). However, for this

reason multiple views are employed that appeal to different (and most often simultaneously active) goal frames.

The *'save the environment'* view displays the total amount of CO_2 that was not emitted as a result of the savings. Trees were used as the bigger units, while the number of kg's of CO_2 saved resembles the absorption capacity of a tree over a year's time. As the smaller unit, small CO_2 clouds were chosen that fill up the tree before being replace with a new one.

The *'have fun while saving energy'* view relates to the game concept employed in enCOMPASS. It depicts the number of points the user has received for his energy savings, while at the same time displaying his/her progress towards the next badge for saving energy, which is one of the thematic areas (see *D2.2 Final requirements*).

The detailed view

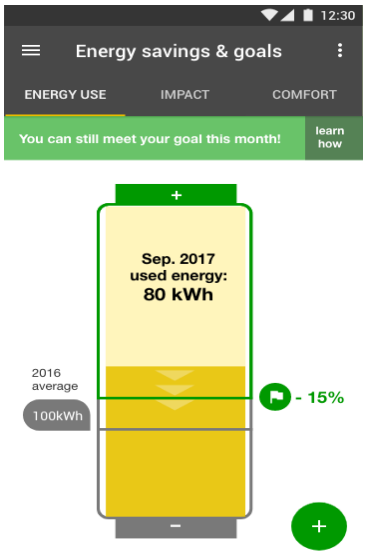
The detailed view is intended for data-affine users who value detailed data-oriented visualisations (Micheel et al., 2014). Such visualizations often display consumption information in e.g. bar or pie charts (Froehlich et al., 2012). In enCOMPASS a bar chart consumption visualization is made available, that allows users to check the evolution of their consumption at different time resolutions (e.g. daily, weekly, monthly), while, in line with recommendations from Karjalainen (2011), it also allows for comparison against the user's own historical average.

3.2 Overview Visualization

We have adjusted the visualizations that we have used in the mock ups. In

Figure 3.1 you can see both these visualizations. The concept of the visualization has remained the same: we use the metaphor of the battery to show how much electricity the users consumed and to monitor how much they have left as opposed to their consumption in the same month in the previous year (this number is shown in the grey box at the bottom of the battery). The idea of this visualization is to allow the users to set a savings goal for themselves – which should range between 10-30% as to what they have consumed last year. The users can set this goal themselves by clicking on the button in the right-hand corner as shown in the Figure 3.2. In the application, the users are provided with tips how they can reduce their energy consumption to reach this goal. The battery then gives an overview and the progression towards this goal. Apart from the visualization of the energy depleting in the battery, there are several indicators in the visualization giving the user feedback on how they are performing: 1) the boxes next to the battery which display the exact amount of kWh of energy that has been used and that is remaining; 2) the color coding; and 3) the message on top of the battery. The user can then use this information to adjust their energy consumption in the remaining time of the month in order to meet the goal. Additionally, the user can also view the weekly progress and see in which of the weeks they have used more electricity and thus adjust their consumption.

Visualization used in the Mock-Ups



Updated version of the visualization

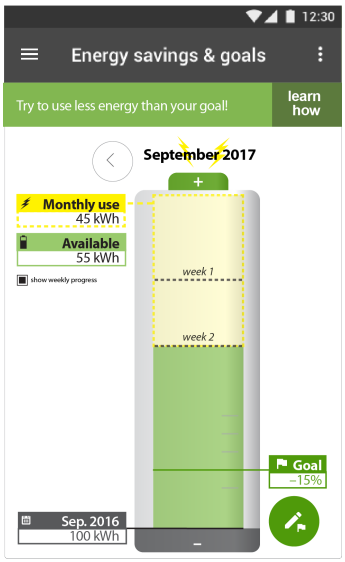


Figure 3.1 Overview Visualizations used in the Mock Ups and to be used in the first version of the application

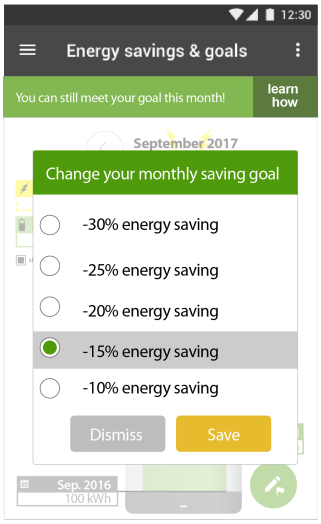
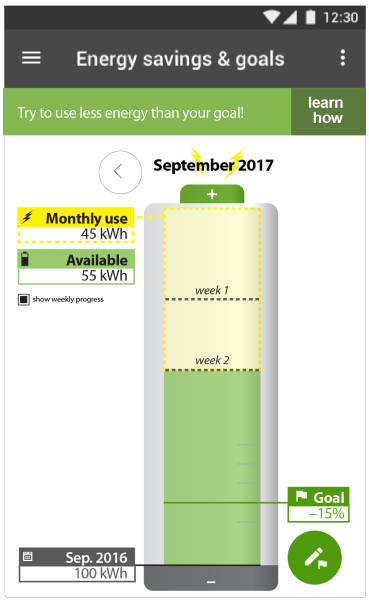


Figure 3.2 Illustration of how the user can set a new goal in the application

One of the main issues that came up during the design of the visualization is the orientation of the battery. One possibility is to "deplete" the battery from the top so that the electricity available is shown at the bottom – this possibility is illustrated in the left-hand side of Figure 3.3 as DIRECTION 1 visualization. Another possibility is to deplete the battery from the bottom – as illustrated in the right-hand side of the Figure 3.3 as DIRECTION 2 visualization. Both of these visualizations have their advantages: in the DIRECTION 1 visualization it seems logical that the battery is depleting from the top, however the goal is on the bottom which can be counter-intuitive. In the DIRECTION 2 visualization the goal is on top – which seems more logical to achieve something. We have tested both these visualizations and the results are presented in Section 5.1. As the results show that users prefer the Direction 1 visualization, for all explanations we will use this view, however the views of the DIRECTION 2 visualization can be seen in Appendix 1 and will be referenced throughout the document.

DIRECTION 1 Visualization



DIRECTION 2 Visualization

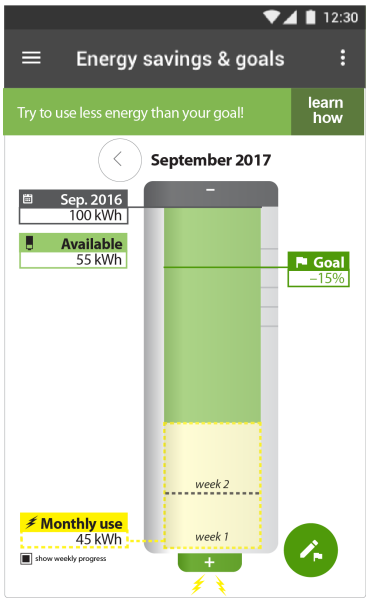
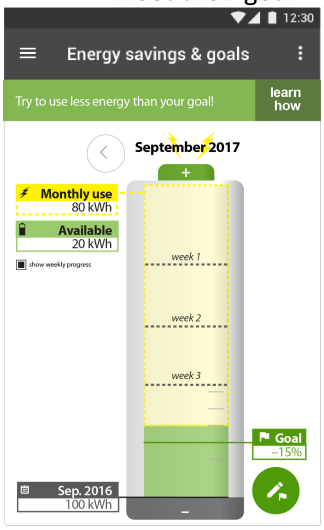


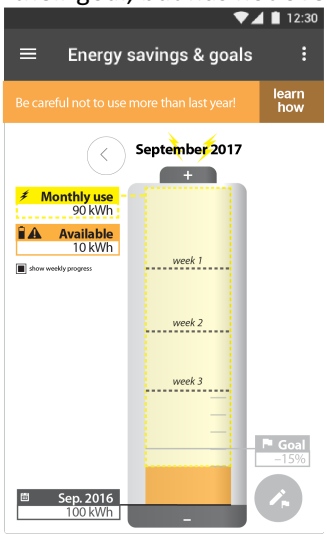
Figure 3.3 DIRECTION 1 vs. DIRECTION 2 presentation of the Overview Visualization

As noted previously, color is used as one of the main ways to communicate to the user how they are doing with respect to meeting their goal in the progress of the month. In Figure 3.4 it is clearly presented that green is used when the user is on the right track to meet their saving goal in the month, orange is used when the user has surpassed their goal, but did not use more than last year, and red is used when the user has done both, surpassed the goal and used more than last year. This "traffic light" analogy has proven itself quite effective in providing feedback to the users. In the same figure, one can also see that other elements reinforcing the message are changing as well: the colors of the boxes with "available energy", the contents and the color of the message on top of the battery and the inability to change the saving goal in the last two screenshots where the savings have not been met.

Green: the user is on track to meet their goal



Orange: the user has surpassed their goal, but has not overused



Red: the user has used more than last year

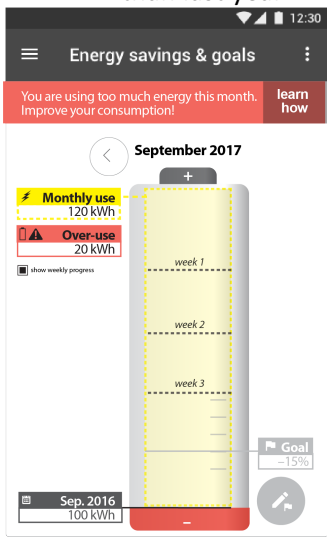


Figure 3.4 Differentiation of states in terms of meeting the savings goal (progress view)[1]

[1] - Here we present the screens for the Direction 1 visualization. The ones for the direction 2 visualization can be seen in Appendix 1

By the same token, we are using the colors to show how the user has performed in the months in the past. The users have the ability to check their performance in the past months by clicking on the arrow next to the month's title. As can be seen in the Figure 3.5, similarly to progress, there are three possible states: green if the user has achieved the savings goal, orange if the user has not achieved the goal, but used less electricity than last year and red if the user has neither achieved the goal nor used less than last year. Additionally, to show that the month has been completed, the boxes with the information about electricity use cannot be edited and are fully colored, alongside an additional feedback about how the user has performed is provided with a smiley next to the amount of electricity the user has used in the month. A "smiling smiley" is given for the case when the user has achieved the goal, the "neutral smiley" is given when the user has not achieved the goal, but used less than the previous year and the "sad smiley" is given for over-use of electricity. So that the users are not confused as to whether the month has passed or is still ongoing, the shades of the color green are used.

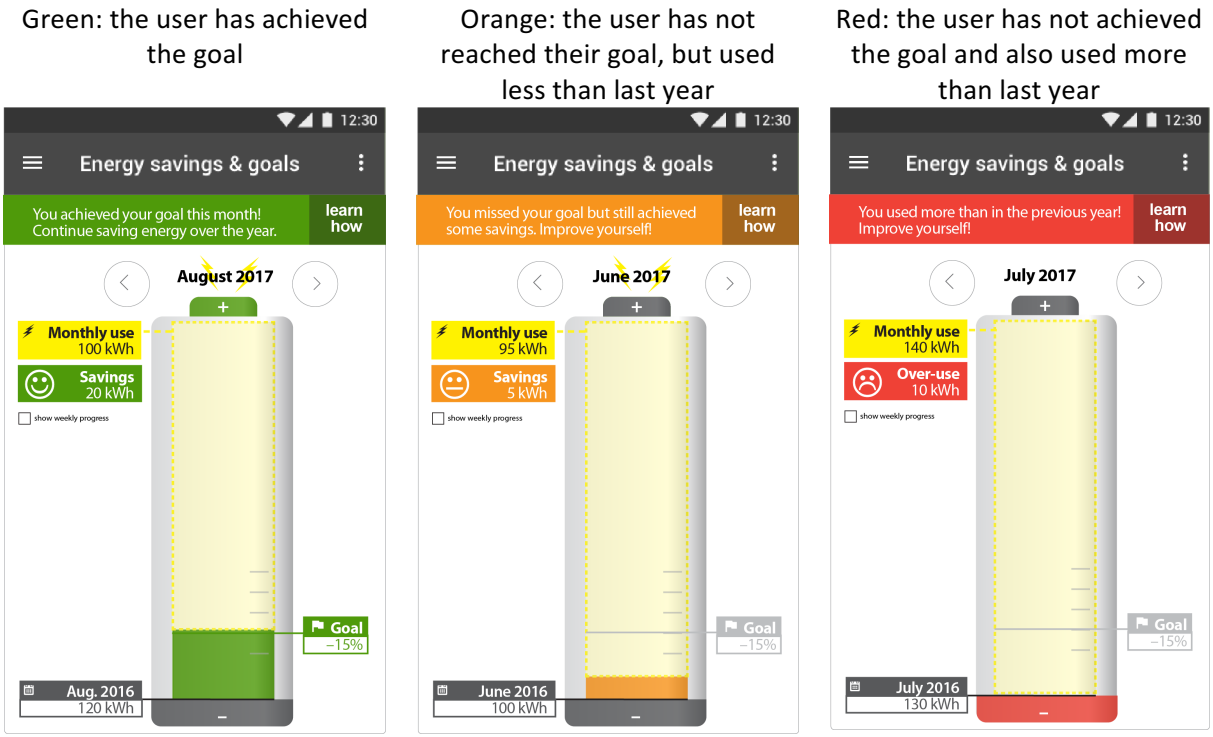


Figure 3.5 Differentiation of states in terms of meeting the savings goal (results view)[2]

For the results of explicitly testing these visualizations and collecting user feedback please refer to Section 5.1.

3.3 Impact Visualization

For the impact visualizations, as discussed in Section 3.1 we have developed three visualizations: the monetary, the environmental and the hedonic visualization. The basic idea behind these visualizations is to display the cumulative savings over the year to the user not only in terms of kWh hours, but also in terms of a more tangible representation of the savings. As the savings are presented cumulatively over the year, we

[2] - Here we present the screens for the Direction 1 visualization. The ones for the direction 2 visualization can be seen in Appendix 1

have to account for quite a large number of potential savings to be presented on the screen. Therefore, we use a two-stage process in all of the impact visualizations as discussed below. For the monetary visualization, as depicted in Figure 3.6, we have used the metaphor of the piggy bank to display the monetary savings that the user can achieve by saving energy. The two-stage concept in this case reflects the coins that can be collected which reflect the actual price of electricity (that will be adjusted for any regional differences). In the example in the Figure 3.6 one kWh of electricity costs 0,28 cents, therefore the user has saved 150kWh or an equivalent of 42 Euro which can be clearly seen by counting the piggy banks without the necessity to consult the table above the image. Once the user has saved 10 Euro, then one piggy bank has filled up. The piggy bank which is still being collected is larger than the piggy banks that have already been collected, so that the user can also see and count the coins contained in it.

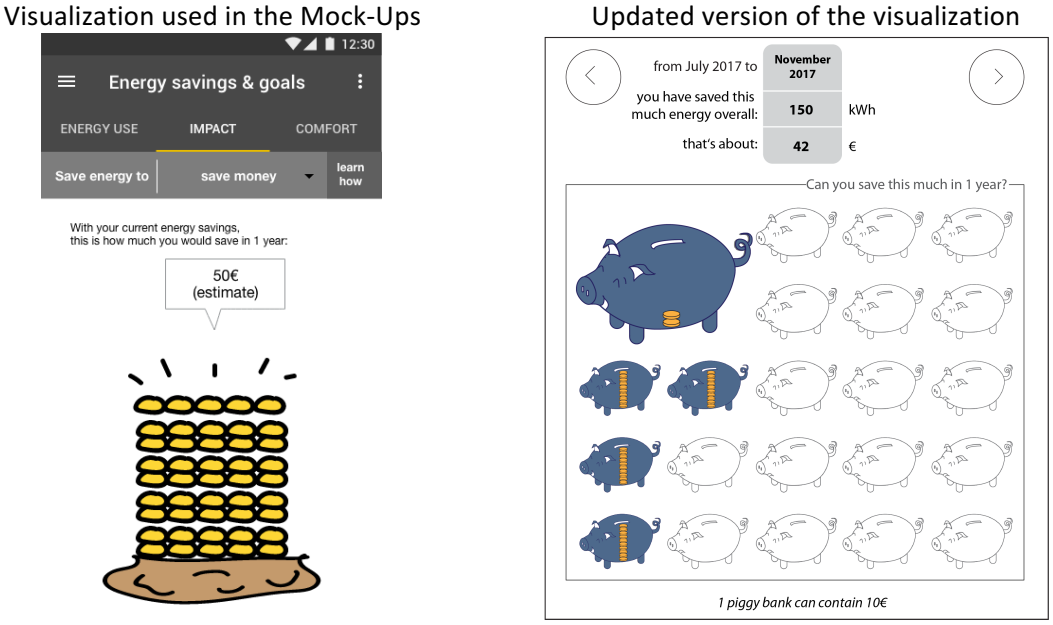


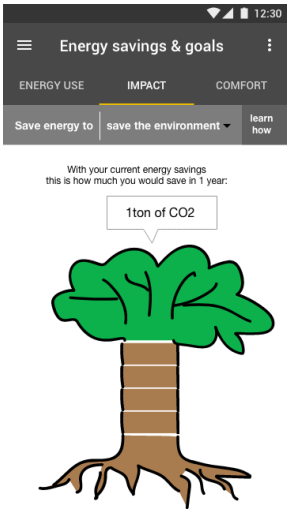
Figure 3.6 The Monetary Impact Visualization

In order to visualize the impact on the environment, depicted in Figure 3.7 we have used the analogy of a tree which does not have to absorb CO_2 if energy is saved. The two-stage concept in this case reflects a CO_2 cloud which would not be emitted into the atmosphere if energy is saved (for 1 kWh of electricity ca. 0,5 kg of CO_2). It is estimated, that a tree can absorb 22 kg of CO_2 over the course of a year, so once this amount of CO_2 is saved, one full tree can be saved (or the work of one tree can be spared). Similarly, due to the abundance of CO_2 elements inside the tree, the current tree which is being saved is larger than the already saved or not yet saved (empty) trees.

The visualization relating to the hedonic motivation is depicted in Figure 3.8 (right). This visualization changed from the initial mockups (foreseeing a puzzle-based visualization) as we refined the concept to better fit the purpose of this visualization and user expectations. The updated version used in the tests reflects in a graphical view the points that one can collect for achieving saving goals and saving energy which are a part of the gamified process in the enCompass application. This relates to the intrinsic sense of satisfaction about the achieved result supported by the gamified mechanics. Designing this visualization, we took the same structural idea as with the other two visualizations: the points are collected into jars, and the layer of "saver status" reflects the purpose to collect points (awarded for achieved energy savings) and make it more engaging for the user. If the users collect a certain number of points, they are awarded a

status of a beginner saver, if they save even more than that, they are intermediate, and the highest type of saver status is advanced. Accordingly, this is also shown in the visualization that mimics the structure of the gamification mechanics for this type of achievement: the users see the required number of jars to be filled to reach the points associated with the given saver status. As they collect the points, the jars are filled one after the other, leading them on their way to the next saver status.

Visualization used in the Mock-Ups



Updated version of the visualization used in the tests

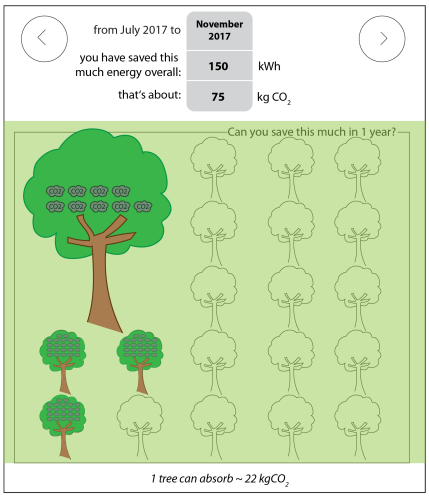
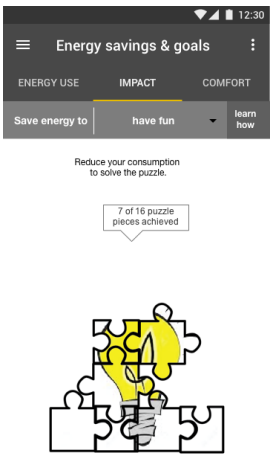


Figure 3.7 Environmental Impact Visualization

Visualization used in the Mock-Ups



Updated version of the visualization used in the tests

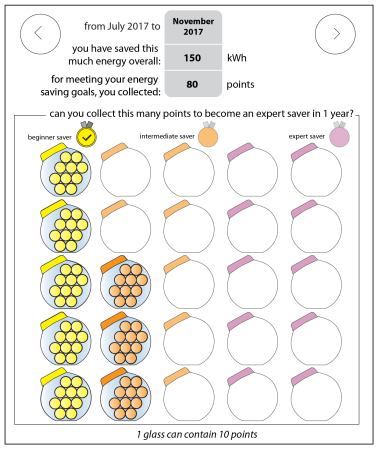


Figure 3.8 The Hedonic Impact Visualization

One of the concepts that spans through the logic of the impact visualizations is the gradual collection of savings and their cumulative presentation to the user as depicted in Figure 3.9. Users are able to click through the periods by clicking on the arrows and the elements that they have saved during this time will be added. Thus, the users get the feeling of the progress they have achieved with electricity savings and the corresponding monetary, environmental or hedonic benefits. In the beginning the screen is empty – the

users have not saved anything. Then the elements start to fill up with a total capacity to be still adjusted depending on how many savings the users can potentially and realistically achieve in their household over the course of the year. The users can consult this screen and determine how they performed in the past as well as set ambitious goals for the future.

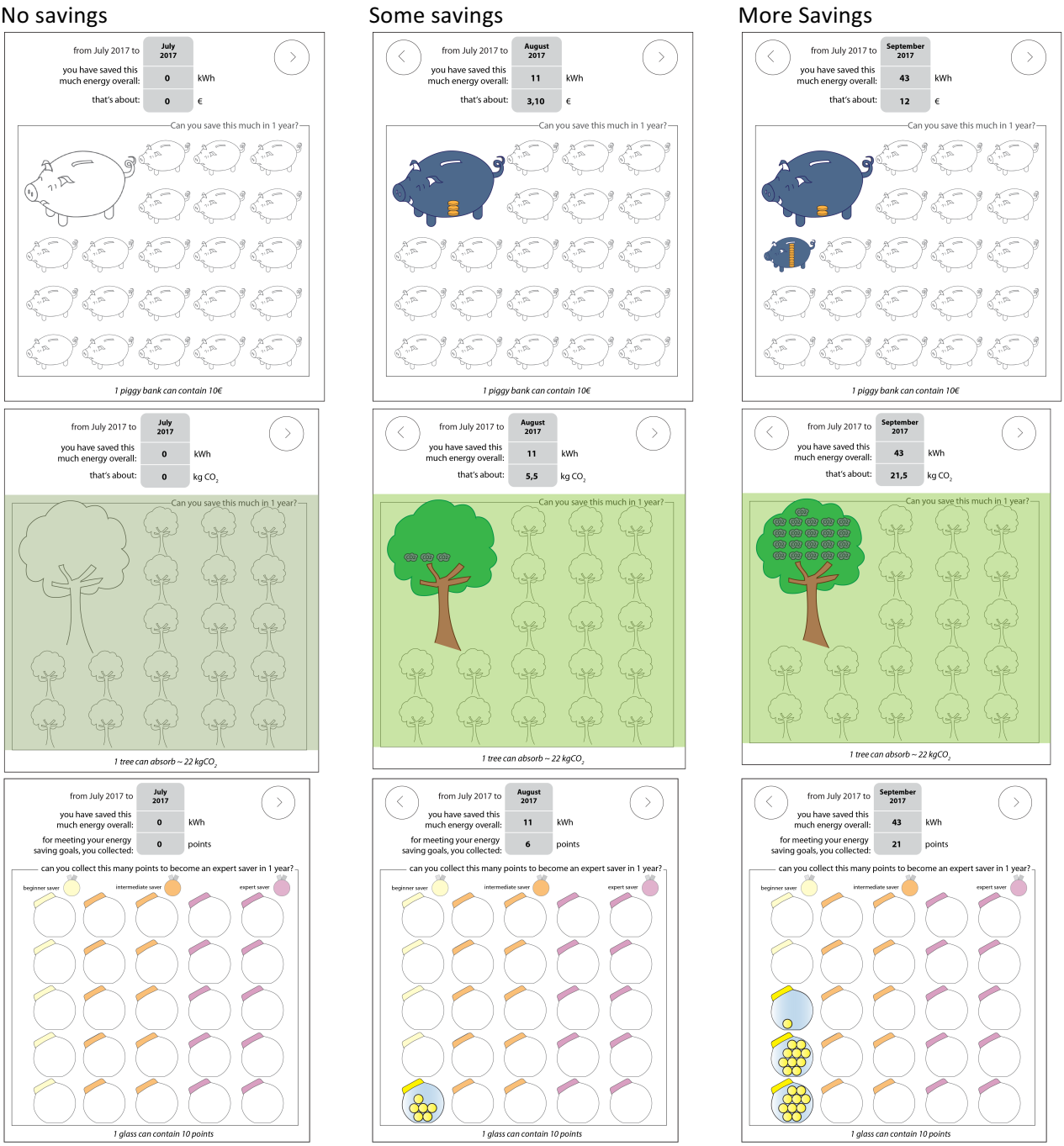


Figure 3.9 Impact Visualizations illustrating the progress of savings over time

For the results of explicitly testing these visualizations and collecting user feedback please refer to Section 5.2.2.

3.4 Detailed Bar Chart Visualization

The bar chart visualization we present to the users is two-fold: we present the data over time and we also present the data per device. The detailed bar chart visualization presented in the left part of Figure 3.10 is

the most commonly used format for consumption feedback, which allows the user to monitor the evolution of their consumption over time. The user's energy consumption is displayed as bar charts. Optionally, users can display a line in the bar charts that display their average consumption level over the last three years. A slider can be used to change the timespan to custom time ranges. In this way, data-affine users can interact with the feedback, and find out opportunities for savings, and distinguish diverging consumption patterns, when e.g. one week's consumption is different from the other. Such close monitoring of consumption induces a deeper insight into one's consumption, and supports the user to reflect on the causes for different consumption levels during different time periods. As such, it stimulates a sense of responsibility for one's own consumption levels.

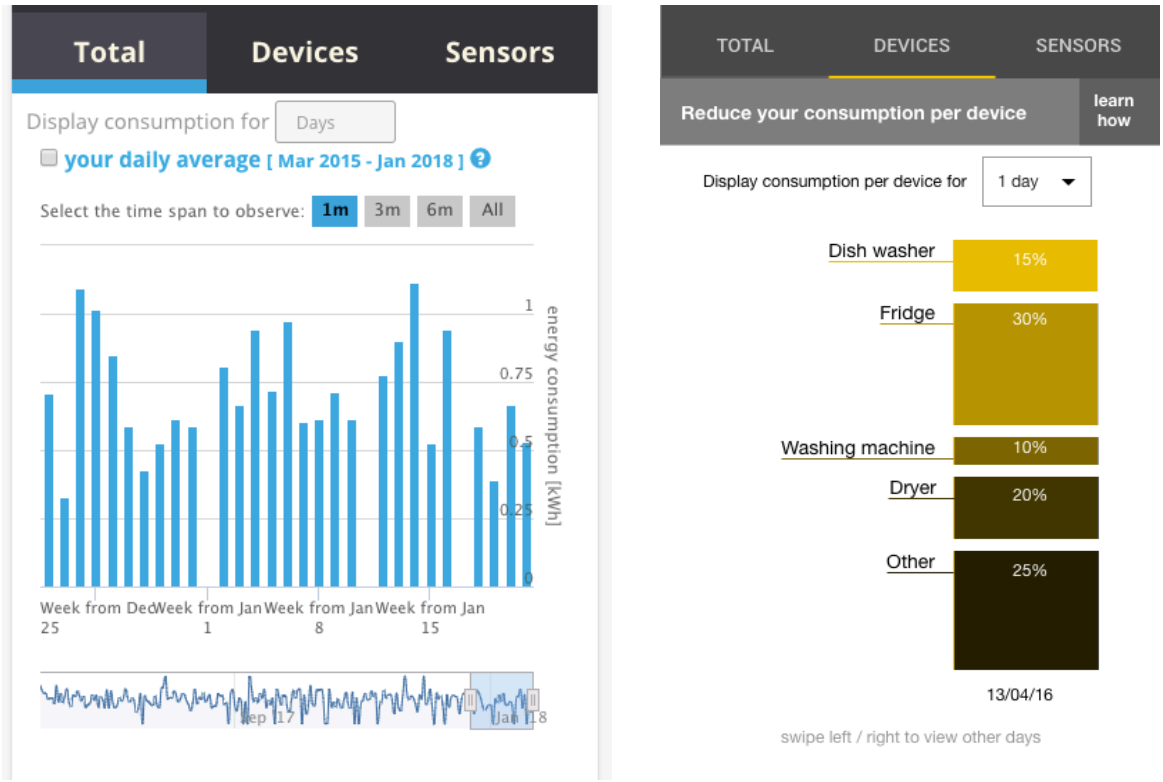


Figure 3.10 The detailed bar chart visualization.

The detailed bar chart visualization presented in the right part of Figure 3.10 is also a common way to present the users with additional information about which devices are using most electricity, and allow the users to adjust the consumption of these devices. For example, if a user notices that the fridge is relatively high in using electricity, then the user can take better note of how the device is used, and by utilizing the tips that are given through the enCOMPASS application adjust it accordingly (e.g. do not leave the fridge open too long). This view also allows the user to see the consumption over a certain amount of time: over a day, week or month. In this way the user can directly estimate in which days there was an overuse on which device and try to change behavior accordingly.

4 First version of integrated gamified awareness apps and game concept

4.1 First Version of the Awareness Application

The technical objective of the first early version of the awareness application is to integrate smart meter readings within the end-user application, while also testing the integration of the gamification engine. As the application is in an early stage of the development process, not all incentives have yet been integrated. However, for the purpose of the test (reported in Section 5.3), all the yet to be integrated functionalities were included as snapshot, so to allow the user to get a tentative impression of the user experience.

As a socio-technical information systems that seeks to influence alter or reinforce attitudes towards energy saving, and ultimately to change energy saving behavior (Oinas-Kukkonen, 2013), the awareness app combines a range of different behavioral change incentives, as feedback alone is not capable of inducing a sustainable change in behavior (Fréjus & Martini, 2016). The incentives and the current state of integration is displayed in Table 4.1.

Table 4.1 Means of the main evaluation items and mean differences between treatments

Consumption feedback Status: detailed consumption view is integrated. For the other consumption views, snapshots were included. Consumption data is retrieved from a working smart meter.	In terms of the norm activation model (Schwartz, 1877), consumption feedback is expected to increase the awareness of consequences, to stimulate ascription of responsibility, and to leverage the impact of communicating norms to the users. The designs proposed in this deliverable were designed to strengthen these behavioral determinants.
Gamification Status: leaderboard, and badges are integrated; actions (e.g. logging in, responding to a tip) yield points; overview of achievements is integrated, just a draft page with rewards.	Gamification, as the addition of game design elements in non-game contexts (Deterding, 2011), was employed to incentivize users who do not yet have a strong intrinsic motivation for energy saving. By appealing to generic human needs such as the desire to compete, or to collect (Reiss, 2014), and by clearly linking game-based rewards (e.g. points, social status, physical rewards) to energy saving achievements, intrinsic motivation is expected to grow with increased exposure. Points, leaderboards, and badges are employed, as well as an overview for the users to keep track of their achievements.
Recommendations and tips for energy saving Status: separate pages are available for tips and recommendations. Feedback options are implemented. Feedback is stored. Integration of R1 Recommender service is scheduled after release of D4.2.	In line with earlier research on the design of energy consumption feedback applications, feedback should best be combined with suggestions for concrete actions (Sundramoortyhy et al., 2010; Micheel et al., 2014; Spagnolli et al., 2011), preferably with feedback options that have the user explicitly express their commitment to carry out the action (e.g. Abrahmse & Steg, 2013). A distinction is made between generic tips, and context-aware recommendations, with the latter being personalized according to the user's current context (e.g. location, activity), and his/her behavioral (consumption) patterns, to optimize the chance of the user responding to the recommendation, and to personalize the recommendations, as tailored recommendations are expected to be more effective.
Comfort feedback Status: comfort level display and	Inducing a deeper insight into the trade-off between visual and thermal comfort and energy saving is one of the objectives of

feedback options are not yet integrated, but included as snapshots. Integration will be completed before R1.	the application. A visualization that temporally compares energy savings against comfort level averages in units understandable to the user is expected to induce such insight. Visualizing comfort for usage by regular users rather than building managers, is a novel area of research.
Notifications to trigger attention Status: integration is scheduled for	Smartphone push notifications are used to continuously trigger the attention of the user, as to avoid the user losing his/her engagement with the app. Push notifications concern all incentives employed in the app (e.g. gamification status updates (e.g. a new badge, or a periodic summary), reactivation messages to users who have become passive, consumption warnings, tips (e.g. tip of the week), and personalized recommendations). Users will be able to adjust the type, frequency, and scheduling of the notifications.

In sum, the first version of the awareness application offers an initial integration of the gamification engine, the detailed consumption view, and the energy saving tips, while the other consumption views are embedded as the snapshots (e.g. an image). For the remaining functionalities (e.g. comfort feedback), the mock-ups are integrated as images.

Gamification is essential to the concept, in that it has the potential of engaging users who are otherwise not intrinsically motivated for saving energy. Sustainability of the approach is fostered by linking gamified incentives and rewards to real-life events associated to energy saving. For completeness, we explain the main gamification concepts employed in this first version of the awareness app (for a more elaborate account of gamification, please see D5.2).

- **Thematic areas:** categories in which the gamification objects (action, badge areas) are grouped and organized. Examples of areas are: education, reputation, socialization and consumption.
- **Credits:** points the user (player) can earn performing actions on the platform.
- **Action:** a rewarded task the customer can perform on the platform (e.g. Read a tip, watch a video). Actions can be repeatable after a given time elapsed or can be set as not repeatable. Actions can be configured as enabled or disabled, setting them as active/inactive.
- **Badge areas:** category used to group badges related to the same topic, but with different levels (e.g. Super Profiler level 1 – Super Profiler level 2, Beginner Saver – Expert Saver).
- **Badges:** virtual recognitions assigned to a user and visible to other users in the community, mostly used to demonstrate consumer status and progress. It is possible to define which actions contribute to achieve a given badge.
- **Goal:** consumption objectives that can be achieved by the user reducing the average consumption.
- **Reward:** physical item that can be redeemed by the customer, using credits earned on the gamification platform. A reward can be configured as available or not.

In Figure 4.1 and Figure 4.2 the main screens of the awareness app are displayed: the main menu, and the home page.

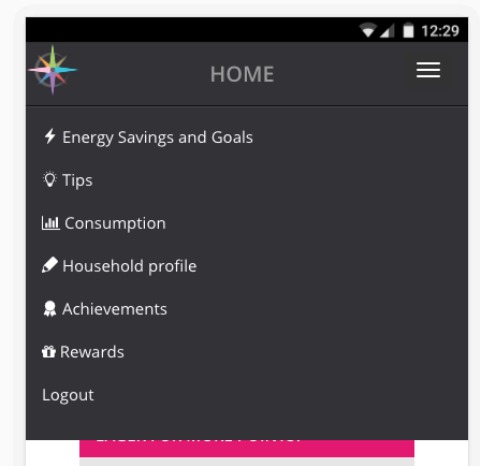


Figure 4.1: Main menu of the Awareness App

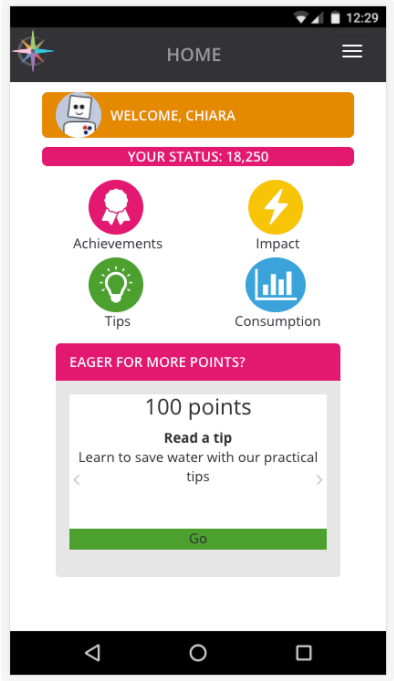


Figure 4.2: Home page of the Awareness App

Below the integrated screenshots of the metaphoric consumption visualizations are shown, which are explained in detail in Section 3.1.

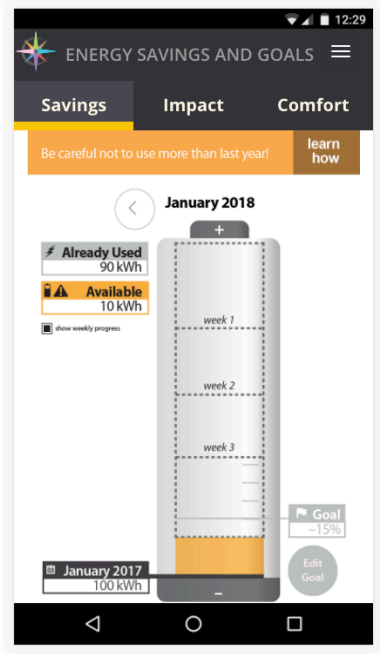


Figure 4.3: Mockup of the Energy saving section

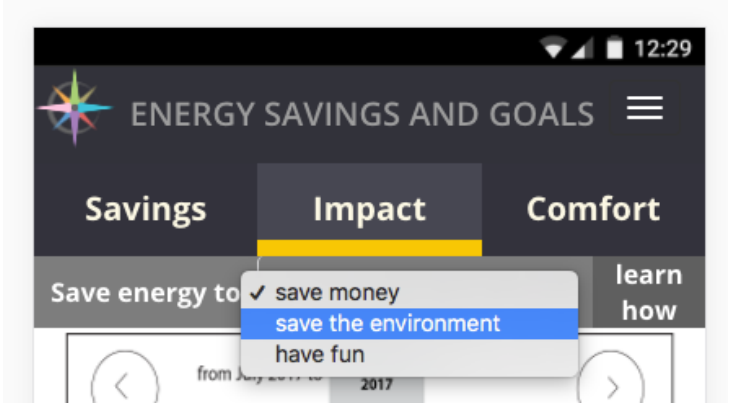


Figure 4.4: Motivation menu

As shown in Figure 4.5, the impact view was embedded as a mock-up, derived from *D2.2 Final requirements*. It represents the inferred thermal and visual comfort level. As it currently is embedded as a screenshot, the button to add comfort feedback is not yet working.

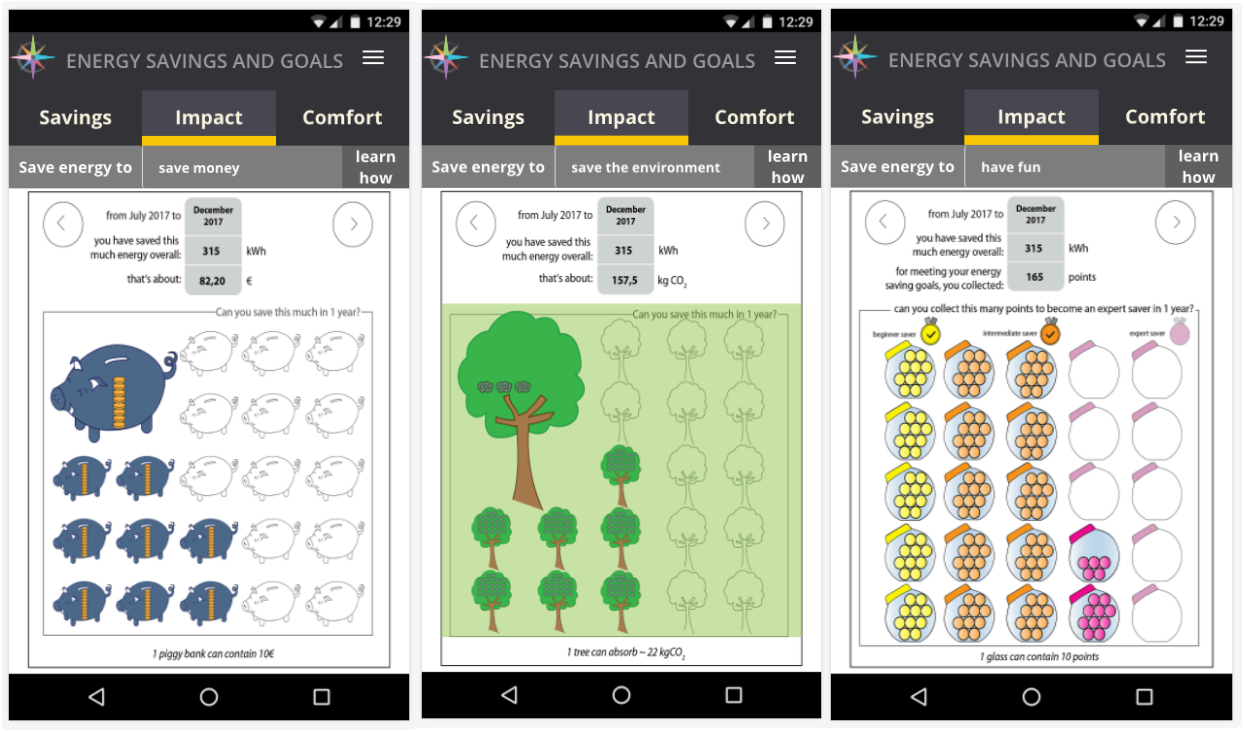


Figure 4.5: Mockups of the impact section

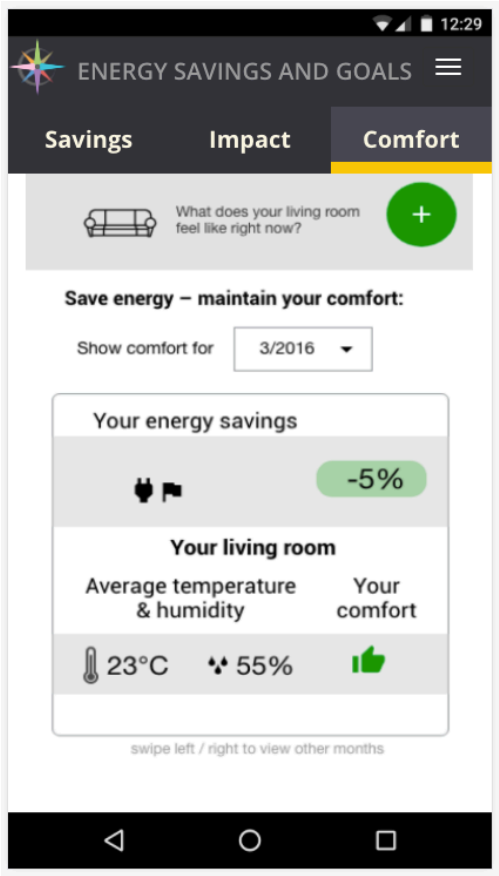


Figure 4.6: Mockup of the comfort section

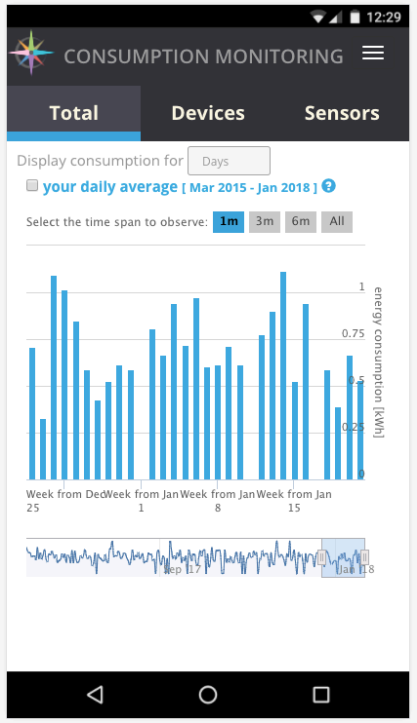


Figure 4.7: Consumption section

The consumption monitoring view contains the detailed consumption overview described in Section 2.4, which allows the user to monitor his/her consumption over time, by choosing different time resolutions, and inspecting individual days, weeks, or months. It also allows the user to compare the current consumption against the historic daily average.

The energy saving tips are implemented, as shown below. The interface allows the user to provide feedback on the tip, which is stored in the central database, and is accessible by the recommender.

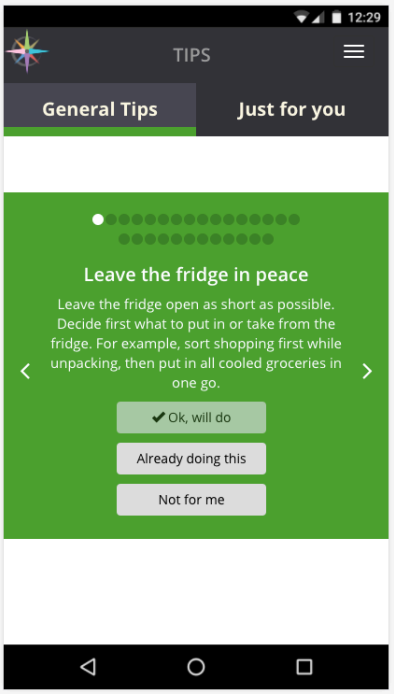


Figure 4.8:Tips section

The screenshots below demonstrate the gamified elements in the awareness app. The achievement page contains the progress and achieved badges for the different thematic areas. The box below the badges allows the user to browse through past actions for which s/he has received points. Two leaderboards are present in the app: the overall leaderboard, and the weekly leaderboard.

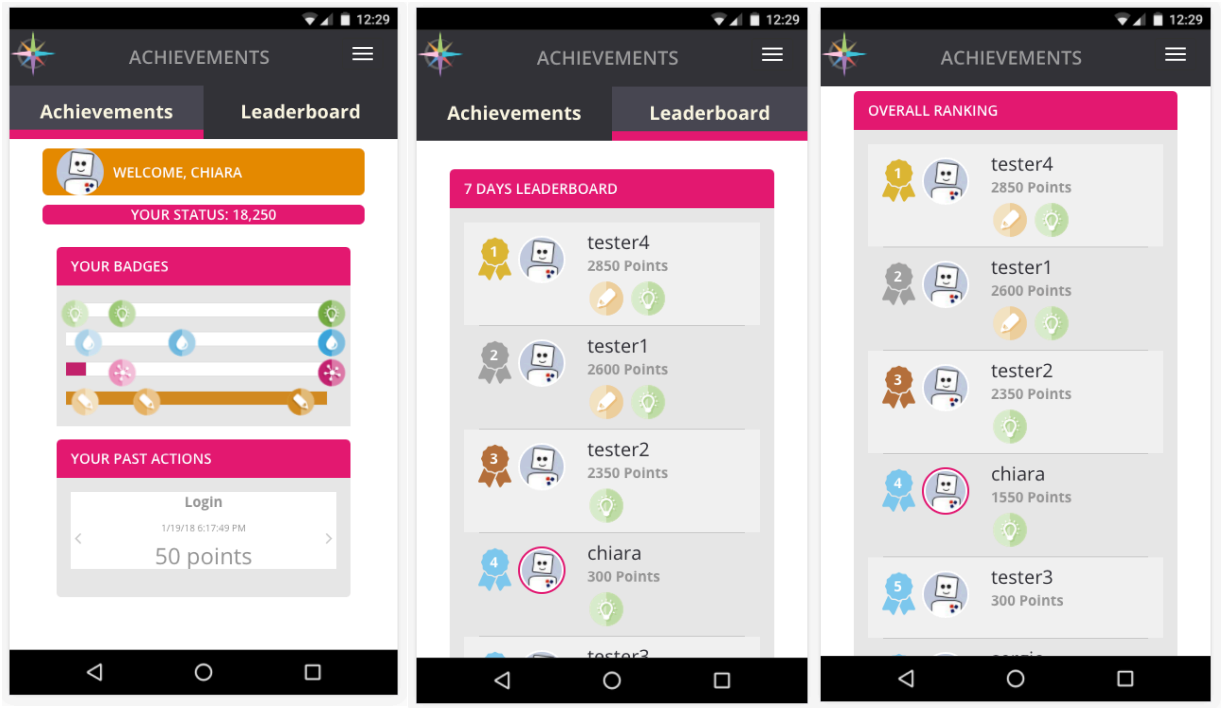


Figure 4.9: Achievement section

The rewards page allows the user to claim rewards when s/he has received enough points to qualify for a reward. The specific rewards available to the pilot users, and the distribution of the rewards will be finalized over the course of the pilot user recruitment process.

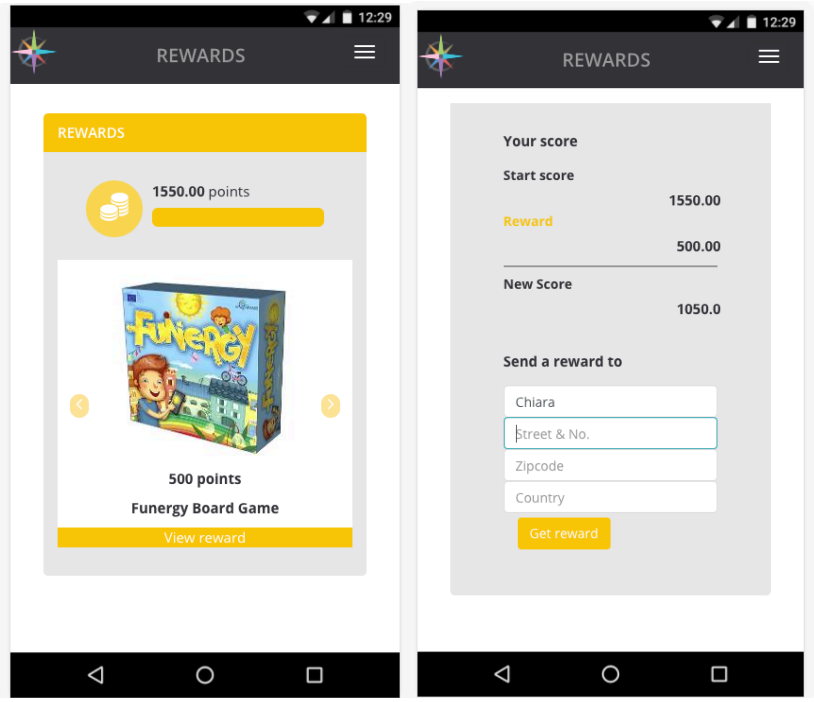


Figure 4.10: Reward section

4.2 INTEGRATION OF BACK-END SUPPORTING GAMIFICATION

The control of the delivery of incentives in the Awareness App is centralized into a component called the Gamification Engine. This component is described in the deliverable D6.2 Platform Architecture and design. In this section we recap only the essential concepts and refer the implementation deliverable for more details. The Gamification Engine "listens" to the actions of the user and transforms them into a variety of rewards, for improving activity and participation.

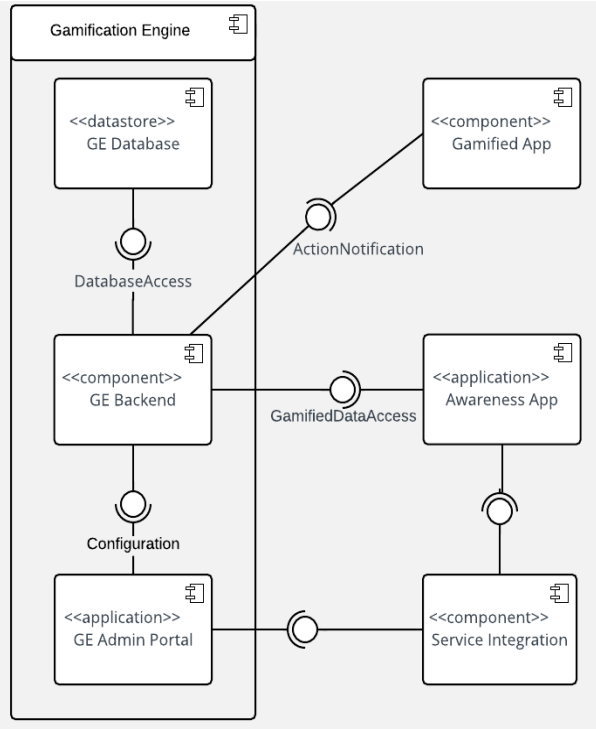


Figure 4.11: Gamification Engine component Diagram

As shown in Figure 4.11, the Gamification Engine is the central component that handles the communication with the Awareness App and other enCOMPASS components and takes care of computing badges, achievements, rewards and all the other gamified features acting as incentives for behavior change.

Its core is the Gamification Engine Backend, which is a parametric rule engine transforming actions into points. All the gamified data are stored in a Central Database, in order to decouple the data from the various energy utilities portals with the one managed by enCOMPASS.

A UI for the Administration of the Gamification engine allows the operator to set the parameters of the gamification rules interpreted by the GE.

The Gamification Engine works as a rule-based engine; it takes inputs and produces outputs as illustrated in Figure 4.11. Its main responsibility is to receive the notification of actions performed by the user and decide if, and to what extent, such actions should be rewarded.

Each **action** is characterized by a set of configuration parameters:

- "oid": <integer>, a unique identifier of the action.

- "name": <string>, a name for the action, meaningful to the user (e.g, “read a tip”).
- "description": <string>, a description of the action for the user.
- "area": <string>, name of the area associated to the action.
- "area_oid": <integer>, id of the area associated to the action.
- "score": <float>, the points earned performing the action.
- "check_time_elapsed" <boolean>, if true, a new instance of the action can be rewarded only after a time interval has elapsed since the last rewarded instance. Incentive models and algorithms Page 38 D4.3 Version 1.2
- "time_elapsed": <float>, null if check_time_elapsed=false: the interval duration for enabling action rewarding.
- "repeatable": <boolean>, if true the action can be repeated multiple times, otherwise it is considered only once.
- "active": <boolean>, if true the action can be considered, otherwise it is ignored.

A **badge** is a virtual reward, characterized by the following configuration properties:

- "id": <integer>, a unique identifier of the badge.
- "title": <string>, a name for the badge, meaningful to the user (e.g., “super saver”)
- "score": <float>, score required to obtain the badge
- "icon": <base64>, icon that represents the achieved badge in the GUI.

Rewards are physical goods that the user can redeem as a result of his actions. They are characterized by the following configuration properties:

- "title": <string>, a unique name meaningful to the user.
- "description": <string>, a description text
- "needed_points": <float>, a number of points for redeeming the reward.
- "available": <boolean>, the availability status of the reward. If false the reward is temporarily out of stock.

4.3 HYBRID DIGITAL-CARD GAME (FUNERGY)

The gamification elements of enCOMPASS are further developed with the use of a Game with a Purpose (GWAP), called FUNERGY. FUNERGY is a card game targeting kids and their families. To exploit the digital attitudes of millennials and recall the digital interactions offered by the enCOMPASS awareness application, the card game is coupled to a digital extension, which has an important role in the game play and further deepens the educational impact of FUNERGY using a quiz-like game mechanics.

4.3.1 FUNERGY card game rules

FUNERGY is a card game targeted to kids. The game concept is to get rid of all my FUNERGY cards faster than the other players to gain the highest ENERGY EFFICIENCY cards and score more points at the end of the game

Components of the FUNERGY card game

The box designed for packaging of the game, shown in Figure 4.12, contains the various elements of the game, 120 cards divided as follows:

- 48 ENERGY EFFICIENCY cards;

- 64 FUNERGY ACTION cards;
- 8 WILDCARDS, which can be branded, e.g., as "enCOMPASS wildcards" .

Figure 4.12: the FUNERGY game box

The ENERGY EFFICIENCY cards

As can be seen in Figure 4.13, these cards are divided into 8 small groups, each of which is marked on the back with a color corresponding to the new European energy scale.

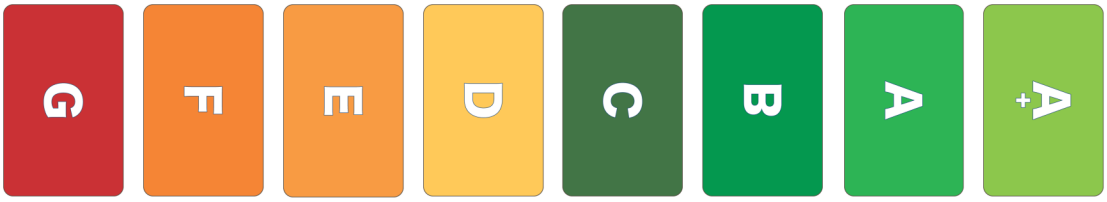


Figure 4.13 Card types in the Funergy Game

On the front, each ENERGY EFFICIENCY card shows the image of a colored piggybank, with a value (see Figure 4.14). The range of values of the ENERGY EFFICIENCY cards of a class increases with the energy efficiency level represented by the class: lowest with G, and progressively increasing. Since the value on the ENERGY EFFICIENCY cards determine total score and thus the winner, the last rounds of the game, with higher values of their ENERGY EFFICIENCY cards, are more valuable to determine the winner.

Figure 4.14 The cards in the Funergy Game

The values on the piggybanks will be summed up at the end of the game to score points.

The FUNERGY cards

All the 72 FUNERGY cards have the same back, depicting the image of the game box.

On the front, the cards show different images and a value that ranges from 1 to 10. The cards are divided in 3 color groups (red, yellow and green). The highest value cards (10) express the most "virtuous" actions, while the red ones show the highest energy consuming actions or devices.

Figure 4.15: Draft of the FUNERGY cards

Among the FUNERGY cards, 8 are marked as "wildcards". These 8 cards have not a value, but a QRcode (see Figure 4.16).

Figure 4.16: Draft of the QRCODE card

Game set-up

• Divide the ENERGY EFFICIENCY cards by colors in bunches and place them in the middle of the table in order to create a sequence starting from the lower efficiency level to the higher efficiency level, as in the example in the previous page.

• Mix all the FUNERGY cards, distribute 5 to each player and leave the remaining ones in the deck, face down on the table

• Draw a FUNERGY card from the deck (if you draw a WILDCARD, place it back at the bottom of the deck and draw again) and place it face up close to the G bunch of cards. This is the "leading card"

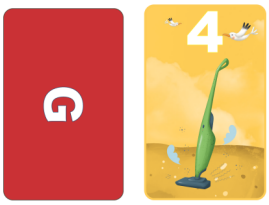


Figure 4.17 Example of a card drawing

• Sort out who will be the first to play.

How to play

The game is divided into multiple rounds, one for each ENERGY EFFICIENCY class.

The player in turn can play one or more of his cards face up over the leading card. In order to do so, the player can play a card whose value is: 1) the same of the leading card; 2) one less or one more than the leading card. The cards just played is the new leading card. For example, if the leading card is a 7, the player can play a 6, a 7, or an 8.

The player can also play a sequence of two or more cards, but in this case the sequence must be ascending or descending. For example, if the leading card was a 7, the player could play the sequence 6, 5, 4 or 8, 9. He could not play, as an example, 7, 8, 9. In any case, after a valid sequence is played, the last card played is the new leading card for next player.

Figure 4.18 Example of a game

EXAMPLE: The player in turn played an 8 and a 9, because the leading card was 7. Now the leading card is 9. Next player can play the same number, a 10, or a 9, starting a descending order sequence. Please note that the images in the example are just placeholders. The graphic of the game is still under development

If the player is unable to play any card, he draws two cards from the deck and then discards one card from his hand, placing the discarded card at the bottom of the deck, face down, without showing it to other players.

The goal for every player is to get rid of all of his FUNERGY CARDS. The one who succeeds wins the round, but be careful**: the last card played must be a green card**. You can't win a round playing a red or a yellow card. If the player in turn plays his last card and this is red or yellow, the round is not over and he must draw two cards from the deck, without discarding anything. After drawing the two cards, his turn is finished, no matter the value and color of the drawn cards.

Wildcards have a special interpretation. The player can attribute these cards any value he wants, but first he must correctly answer a question, using the QRcode and the FUNERGY GAME APP.

If the answer is right, the player can choose any value and declare it aloud. This value can let him play a sequence of cards.

In example: the leading card is a 7 and the player still has three cards in his hand, a 5 a 4 and the WILDCARD. He decides to play the WILDCARD and answers the question correctly. He then decides that the WILDCARD value is 6 and this lets him play also the 5 and the 4 in sequence, winning the round!

If the answer is wrong, the player to the left of the one in turn can use the WILDCARD as he likes, giving it any value he wants.

A player cannot finish the round playing a WILDCARD if he did not answer the question correctly. In case of a wrong answer, he must draw 2 cards from the deck and the game goes on.

End of the round

When a player plays the last card from his hand, the round is over. The winner of the round takes the bunch of ENERGY EFFICIENCY cards and from this he takes for himself the highest value card, distributing randomly the remaining cards, one to each player, and putting aside the others. Players can look at the cards received, but then these cards are left face down on the table

Each player takes from the deck as many FUNERGY cards as needed to have again 5 in his hand, so he draws a new leading card from the deck putting it alongside the new ENERGY EFFICIENCY bunch for a new round.

End of the game

At the end of the last round, corresponding to the last ENERGY EFFICIENCY bunch of cards, the players with the highest total of ENERGY EFFICIENCY cards in his hands is the winner of the game.

4.3.2 FUNERGY Digital game extension

The FUNERGY card game comes with a digital extension, which can be used to further animate the game play while injecting into it further elements of energy efficiency awareness.

The essential requirements for building the digital game extensions are as follows:

- The game must be playable by the following categories of players: players of the FUNERGY card game, users of the enCOMPASS platform and casual players.
- The game play must be simple and quick, not to interrupt too much the course of the FUNERGY card game.
- The game play must adopt a well know game pattern, so to avoid any learning curve.
- The game play must convey some energy awareness content, without jeopardizing the playability.

Based on such requirements, the choice has focused on a simple binary quiz game, in which much of the fun of the game is obtained by the variety and curiosity of the questions, which must be designed to

provide a good mix of purely educational and entertainment content. Furthermore, questions have three level of difficulty (low, medium and high), which may serve the purpose of increasing the level of difficulty during a session, so to make the challenge more difficult as the FUNERGY game play proceeds.

4.3.2.1 Example of game screens

In the following, we briefly illustrate the main screens of the first version of the FUNERGY digital extension game. At the game start, home page shows three buttons, as shown in Figure 4.19.

Figure 4.19: Initial screen of the game

The Single Player button starts a game play. The Decode a card button starts the scanning of the QR code on a FUNERGY card. The setting button opens the game settings editor screen. Figure 4.20 shows the scanning of a FUNERGY card with the digital extension game.

Figure 4.20: scanning a FUNERGY card

Figure 4.21 shows the question answering screen, which contains the text of the question, possibly enriched with multimedia content, and the buttons to provide the answer.

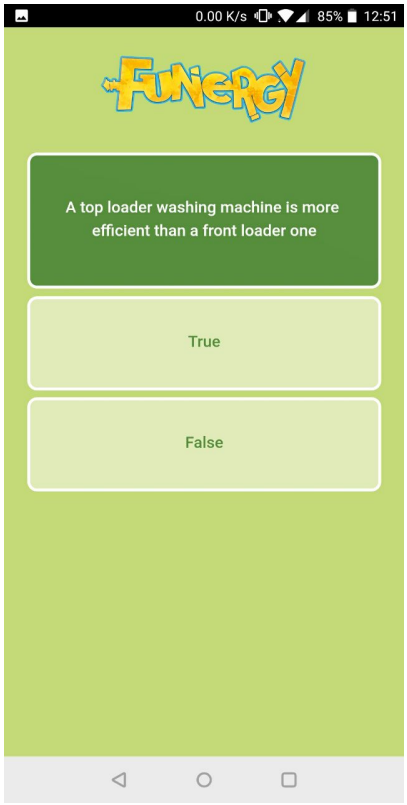


Figure 4.21: question answering screen

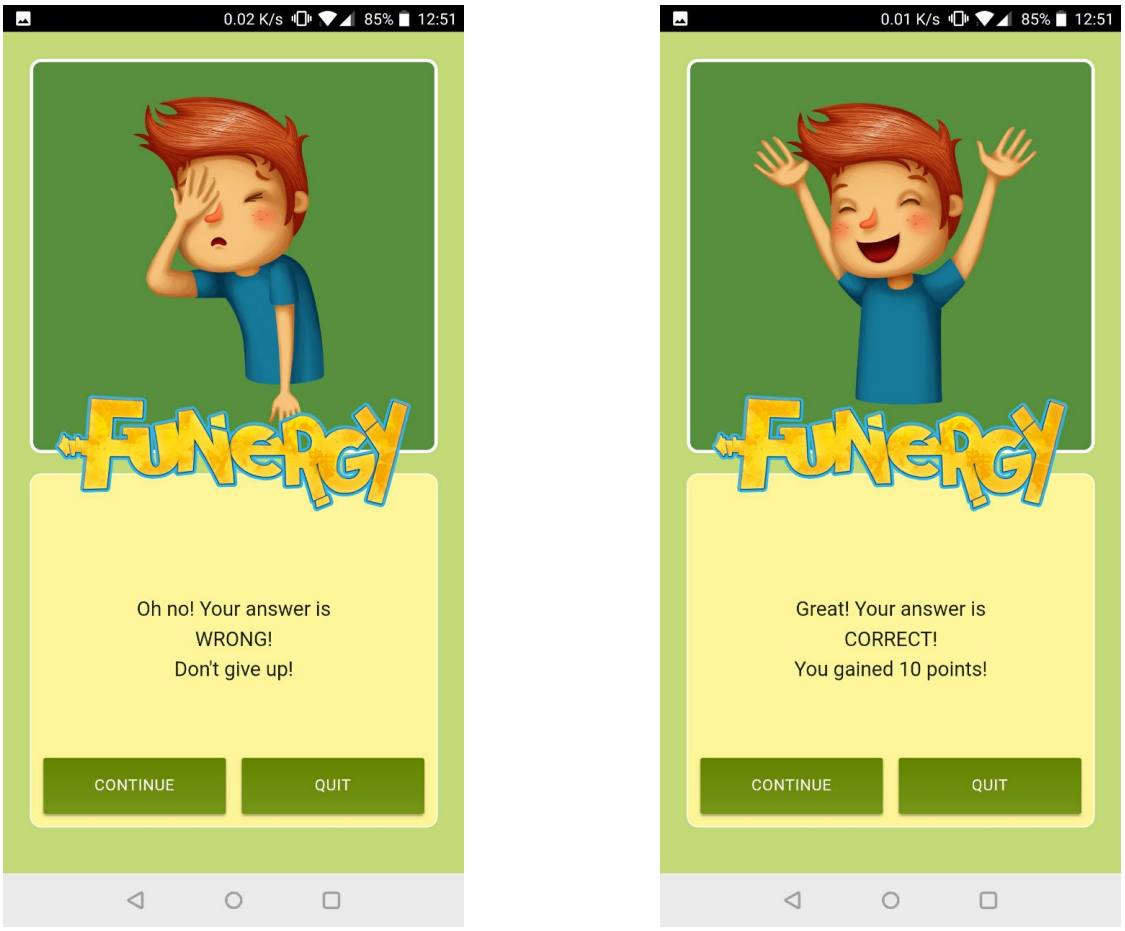


Figure 4.22: result screens after submitting the answer

4.3.2.2 First examples of questions

The following list showcases the first questions collected. Questions are collected, edited and translated with the system available at the URL: http://funergy.ifmledit.org/funergy (password protected).

Table 4.2 First questions collected for Funergy digital game extension

Number	title	difficulty
1	What is the energy measurement unit?	1
3	Using fans instead of air conditioning can reduce electricity consumption a lot	1
4	A top loader washing machine is more efficient than a front loader one	2
5	Leaving electrical devices, such as the TV, on stand-by does not affect the electricity bill	1
6	What percentage of a home's energy use is consumed by water heater?	2

Number	title	difficulty
7	Are instantaneous hot water heaters more efficient than tank water heaters?	2
8	What does LED stand for?	2
9	What is a cool roof?	1
10	What kind of heat pump uses the constant temperature just below the ground to heat or cool homes?	1
11	What are phantom loads?	1
12	How much of the energy in an average fossil fuel-burning power plant is lost as waste heat?	2
13	How much of energy in the fuel for a car is used to move it ?	2
14	Which type of light bulb is the most efficient at converting energy into light?	1
15	When you first turn on an air conditioner, it's best to set the temperature as low as possible to cool faster.	1
16	What does the Jevons' Paradox say?	2
17	In 2016 China was the country that consumed most energy, which country came 2nd?	1
18	In 2016, which country had more than 60% of electricity production from renewable sources?	2
19	Using a screen saver saves energy	2
20	What is the voltage in a typical household electrical circuit in Europe?	1
21	What is the unit of electric current?	1
22	What is the electric current?	2
23	What is the voltage?	1
24	Which movie (starring electricity) was nominated for the Oscar?	1

Number	title	difficulty
25	Which movie featured Nicolas Tesla, the famous inventor of alternating-current electrical system?	2
26	Which American poet wrote the poem "I sing the body electric"?	2
27	Who wrote the science-fiction novel "Do Androids Dream of Electric Sheep"?	2
28	What was the first electronic musical instrument?	3
29	Who discovered that lightnings are made of electricity?	1
30	What material was used in the first bulb lamps?	2
31	What is a greenhouse gas?	1
32	Which of these types of bulbs are the most efficient?	2
33	Do compact fluorescent (CFL) bulbs contain gas?	2
34	Do halogen bulbs contain gas?	2
35	Do halogen bulbs contain a tungsten filament?	2
36	Do halogen bulbs light up instantly?	2
37	Can you use light dimming switch for compact fluorescent light bulbs?	3
38	When there is no light bulb in the the socket but the light switch is on, is electricity being used?	1
39	Does a dimmed light bulb functionally use less electricity?	1
40	Why do incandescent bulbs glow dimmer with use?	2
41	Which form of energy are ultimately derived from solar energy?	2
42	Every year, wind turbines kill more birds than domestic cats do	2
43	What uses more energy? Six hours of laptop use or making coffee?	2
44	EU energy labels use letters to describe an appliance's energy consumption, which letter denotes the MOST energy efficient types?	1

Number	title	difficulty
45	What is the recommended setting for a hot water cylinder’s thermostat?	2
46	We waste more power turning lights on and off than just leaving them on	2

5 USER TESTING AND FEEDBACK COLLECTION

5.1 RESULTS FROM OVERVIEW VISUALIZATION TESTS

5.1.1 Methodology

In order to test the understandability of the overview visualization and decide which of the overview visualizations (DIRECTION 1 or DIRECTION 2) to display to the users as discussed in Section 3.2, we conducted a survey. The survey was developed in an iterative process and aimed to test several things: 1) the conceptual understandability of the overview visualization in both versions; 2) the understandability of specific UI elements; 3) the attention check about the context of the visualization; and 4) the evaluation of the visualization along the usability dimensions. The questions related to the visualizations, and the users were presented with the screenshots of the visualization to which the questions referred. Sometimes it was difficult to work with static images as the application usage assumes certain dynamics. To illustrate dynamics, we have utilized animated gifs and provided instructions to the users to imagine themselves using this application in a specific context.

In total, the users were presented with 40 questions. Most of the understandability questions were formulated by authors and adapted to the energy saving context, whereas the evaluation dimension has been taken from two sources – two items each measuring the hedonic (quality stimulation) and pragmatic dimensions of interactive products usability (Hassenzahl 2004) and one item each measuring the hedonic and utilitarian dimensions of consumer attitudes to account for the preferences for a certain visualization (Voss et al. 2003). Different format of questions was used: True/False statements, answer option choices, open ended questions. For open ended questions, we set a minimum limit of characters that the users have to write to make sure we get extensive responses which can be later used for qualitative analysis. Sometimes, if participants did not give a correct answer to a closed question, they were prompted to explain their answer in the open-ended form. This set up gave us the ability to collect quite a large amount of both quantitative and qualitative data to understand where the main understandability issues lie and how they can be corrected.

The set up was similar to a between-subject experimental setting. We had two separate groups of participants filling out the survey: for the DIRECTION 1 and for the DIRECTION 2 visualization. The survey questions used in both main tests were identical, only the visualizations differed. The survey was conducted through Qualtrics customer interaction analysis software. The full set of questions that were used in the survey can be found in Appendix 2.

The data was collected on Amazon Mechanical Turk. This is a platform that allows people to do micro tasks for small rewards. It works in the following way: the requester posts a task and required qualifications for this task and the number of tasks that needs to be completed, the workers self-select themselves to tasks they want to complete, after they are done the requester verifies if the works meets quality criteria and approves (or disapproves) it, and if the work meets quality criteria, the workers subsequently get paid a specified amount. The platform is increasingly used by researchers as well as organizational managers to collect data about products, opinions or to carry out small tasks such as transcriptions.

For all of the tests we have conducted, the minimum requirements to participate in the test were that the user is from Europe (to make sure that the population matches the future pilot's population), that the approval rate for the tasks they have completed through the platform is higher than 95% (to make sure that the quality of the delivered work is good), and that they did not take part in our surveys previously (as this might cause response bias). It was estimated that the user would take 25 minutes to complete the

overview visualization survey, for which they were rewarded with $1.25 which although low if extrapolated to an hourly wage, is a comparatively high reward for this platform. An example of the task description which was provided to the workers can be found in Appendix 3. The workers were allotted 1 hour to complete the survey, as some unforeseen circumstances might prevent them to complete in 25 minutes and then they would not get paid.

To test an overview visualization, we have conducted four tests. First, to gain familiarity with the platform and to validate our testing methodology, two pretests with 10 participants each were conducted. The pretests used the DIRECTION 1 visualization. The survey and the visualizations were optimized using the pre-test results until deemed satisfactory. The final version of the survey can be found in Appendix 2.

In the next step, two main tests were conducted to test understandability of DIRECTION 1 and DIRECTION 2 visualizations, respectively. The DIRECTION 1 visualization test was posted on December 18th, 2017 and the data was collected by December 20th, 2017. 31 participants took part in the test and they took on average 27 minutes to complete the survey. All of the assignments were of adequate quality and were approved. The DIRECTION 2 visualization test was posted on December 20th and the data was collected by December 24th. 30 participants took part in the test and took on average 26 minutes to complete. All of the assignments, again, were of adequate quality and were approved. In fact, in many instances the users provided explicit feedback and wrote many more sentences that were required from them. This fact, combined with a relatively fast process of data collection, provides support for choosing Mechanical Turk as the platform to collect the user feedback.

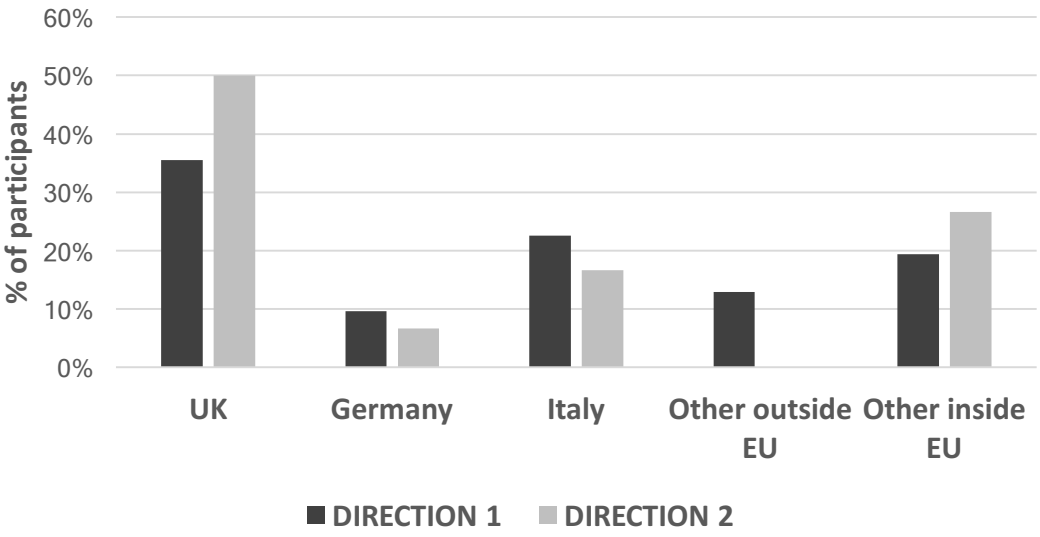


Figure 5.1 Country of Origin of respondents in the test evaluating Overview Visualizations

In the following, we will present the demographics of the users from the two tests we have conducted, the DIRECTION 1 test and the DIRECTION 2 test. From the Figure 5.1 we can see that a large part of the participants came from the UK, the next country with most participants was Italy, followed by Germany. Other respondents were from inside the European Union: France, Poland, Romania, Spain, Croatia, Portugal, Greece, etc. and some respondents from outside of the EU. Although most of the workers on MTurk are from the United States, we specifically excluded non-European countries from our sample (but allowing for non-EU countries), so that the overall context of the survey participants is relatable to the future participants in the enCOMPASS pilots.

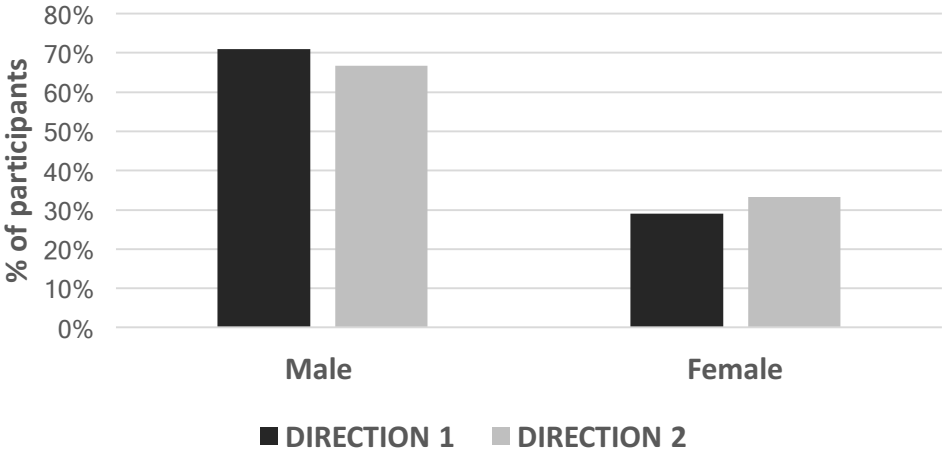


Figure 5.2 Gender of respondents in the test

From the Figure 5.2 we can see that more male than female respondents took part in our test – which probably is explained by the fact that male users have more interest in energy related issues. The Pearson Chi-Square test $\chi(1) = 0.132$; $p = .717$ shows that there is no statistically significant association between Gender and Treatment to which participants were assigned.

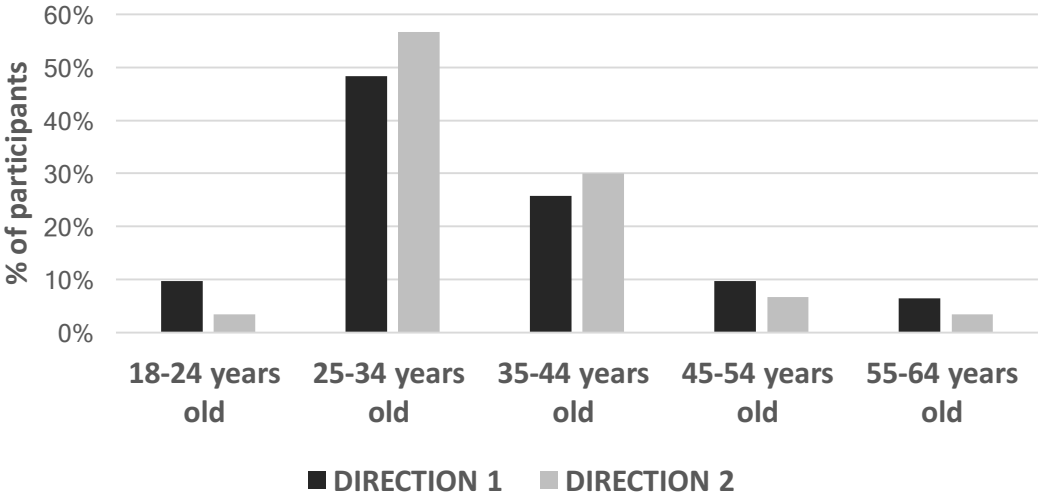


Figure 5.3 Age of respondents in the test evaluating Overview Visualizations

Figure 5.3 reveals that most of the participants taking part in our test were rather young – between 25-34 years of age. This can be explained by the fact that the workforce on MTurk is technology affine, and therefore tends to be rather young. However, we also see users from different age groups, adding to the variety of our sample. The independent samples t-test conducted in SPSS reveals that there are no notable differences between the means of the age in direction 1 and direction 2 groups $t(56.96) = 0.204$; $p = .839$.

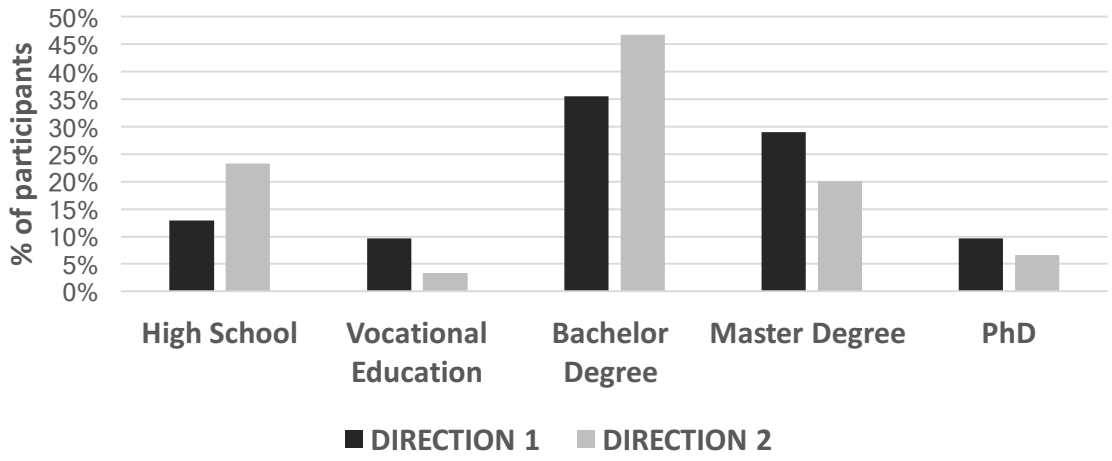


Figure 5.4 Educational Level of respondents evaluating Overview Visualizations

Figure 5.4 reveals that the educational level of participants seems to be rather high – 74% of participants in the direction 1 treatment group and 73% in the direction 2 group have at least a Bachelor degree or higher. However, no notable differences between the Direction 1 and Direction 2 treatment groups were found concerning education t(58.95) = 0.748; p = .458.

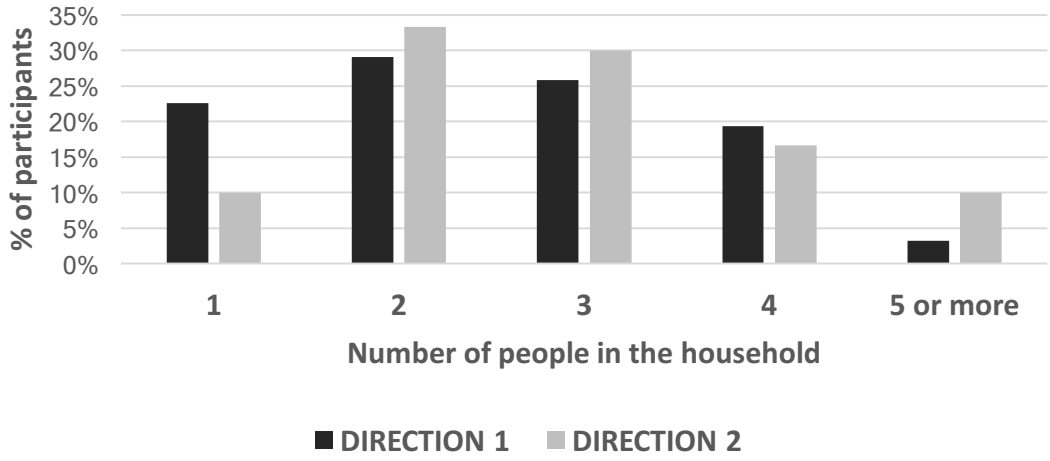


Figure 5.5 Number of people in the household of the participants evaluating Overview Visualizations.

Figure 5.5 reveals that the number of people in the household is on average 2,52 people for the direction 1 and 2,83 for the direction 2 treatment. However, as with previous demographics, no notable differences between the Direction 1 and Direction 2 groups were found concerning household size t(58.94) = -1.078; p = .285. Taking together these results, we can state that the differences between the treatment groups that we might discover in the next section 4.1.2. will not be due to the differences implicit in the characteristics of the sample.

We have obtained a vast amount of data through our tests which we have analyzed qualitatively in the tool called Dedoose and quantitatively with SPSS. The results of these tests will be presented in the next section.

5.1.2 Test results and user feedback

Main aims of the overview visualization test were (in order of importance): 1) to decide which visualization is better understandable to the participants, Direction 1 or Direction 2, 2) check the understandability of the concept behind the battery visualizing energy savings; 3) obtain user's opinions about the visual appeal of the visualizations and their ability to induce them to save energy; 4) obtain qualitative feedback about what the participants liked and did not like about the visualizations; and 5) check understandability and usability of several UI elements. We will thus present the results in this order as well.

The first goal is to decide which visualization is preferred by users: Direction 1 or Direction 2. For this purpose, we included questions aiming to measure different dimensions of their attitudes towards the visualization as well as their overall evaluation of the visualization. To measure attitudes, we used two approaches, one suggested by Hassenzahl (2004) that identifies two dimensions: the hedonic quality (measured by items such as clear-confusing and lame-exciting) and the pragmatic quality of interactive products (measured by challenging-easy and complicated-simple), complemented by the measurement of beauty (measured by beautiful-ugly). Another approach is the standard marketing approach (Voss et al. 2003) to identify the hedonic (by the item fun – not fun) and utilitarian (effective – ineffective) aspects of consumer's attitudes to products. The questions 4, 21 and 24 in Appendix 2 were asked on a 5pt scale in form of a semantic differential, and the average results given by participants in both the direction 1 and the direction 2 treatment can be seen in the Figure 5.6. We observe that the direction 1 visualization seems to be slightly clearer and simpler than the direction 2 visualization, whereas effectiveness and beauty are the same between the two visualizations presented. The mean difference tests, presented in Appendix 5, did not reveal any significant differences in the evaluation of these items between the two treatment groups.

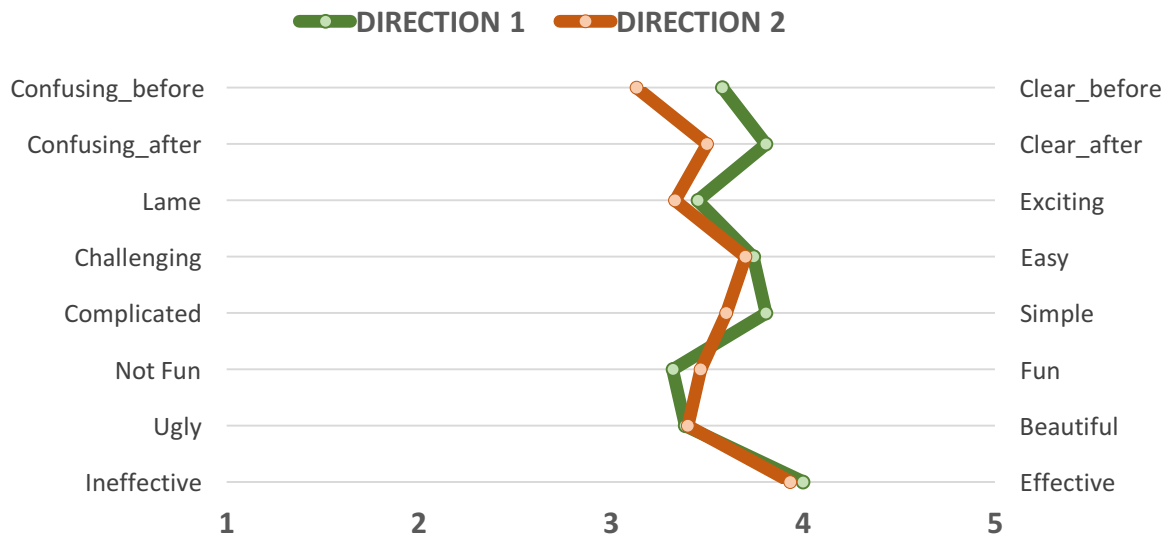


Figure 5.6 Evaluation the two treatments of the visualization along the dimensions of Hassezahl and Voss.

The reader might be wondering about the first two items in Figure 5.6 relating to the confusing-clear evaluation of the visualization. The first evaluation was collected in the beginning of the survey right after the users were presented with the first screen (refer to question 4 in Appendix 2) and the second one - in the end of the survey together with the general evaluation of the visualization (refer to the elements of the question 21 in Appendix 2). This was done to see if the visualization becomes clearer once the users have gained some experience with it and to see which issues are hardest to understand. As we can see from the first two items presented in Figure 5.6, as expected the visualization becomes slightly clearer after they

have gained experience with it, in both treatments this evaluation rises by about half a point. We think that this item is especially important when deciding which visualization to choose, as understanding which visualization was clearer in the beginning is important. Figure 5.7 shows the distribution of the evaluations in both treatments, and we can observe that in the Direction 1 treatment more participants find the visualization clear or very clear (65%) than in the Direction 2 treatment (53%). Moreover, in the Direction 1 treatment the visualization becomes "very clear" to more participants (16%) than in the Direction 2 treatment (7%). Therefore, this provides us with another reason for the choice of the Direction 1 visualization.

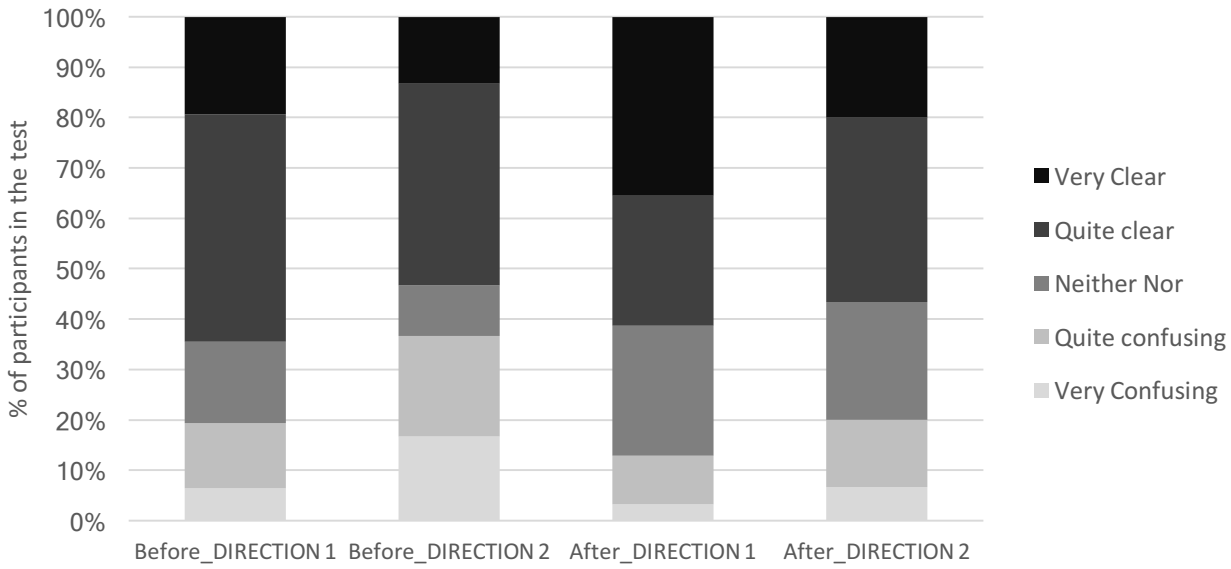


Figure 5.7 Evaluation of the dimension (clear-confusing) in the beginning and in the end of the survey in Direction 1 and Direction 2 treatments

In Table 5.1 we can look closely at the means and mean differences between the treatments for the three main evaluations – hedonic quality, pragmatic quality[3] and the overall evaluation. Judging by this table we can see that although there are no significant differences in the means between the direction 2 and direction 1 visualizations, the means themselves are slightly higher in the direction 1 than the direction 2 treatment for all of these dimensions. The fact that these differences are not significant might be due to the fact that we are running the tests on a small sample size (30 participants in the group).

[3] - Hedonic quality was obtained by summing up the values of clear-confusing and lame-exciting and the pragmatic quality by summing the scores of challenging-easy and complicated-simple

Table 5.1 Means of the main evaluation items and mean differences between treatments

	treatment	N	Mean	Std. Deviation	Std. Error Mean	Mean difference test
Hedonic quality	direction 1	31	7.19	1.493	.268	t(59) = .362, p= .718
	direction 2	30	7.03	1.938	.354	
Pragmatic quality	direction 1	31	7.61	2.076	.373	t(59) = .913, p = .365
	direction 2	30	7.10	2.310	.422	
Overall evaluation	direction 1	31	7.42	1.478	.265	t(59) = 1.011, p = .316
	direction 2	30	6.97	1.991	.364	

The main aim behind the visualization is that it should induce the users to save energy once they have seen it. Therefore, questions 19 and 20 in Appendix 2 are measuring the motivation of users to save energy once they are presented with the visualization. In Table 5.2 we can see the means and mean difference tests of the answers to these questions by the participants in both groups. Again, we notice that although there are no significant differences between these groups, the means do differ and indicate that people in the direction 1 treatment were slightly more motivated both to save energy and to reach their savings goal.

Table 5.2 Means of the motivation items and mean differences between treatments

	treatment	N	Mean	Std. Deviation	Std. Error Mean	Mean difference test
Motivation to reach the savings goal	direction 1	31	4.32	.702	.126	t(59) = .895, p = .375
	direction 2	30	4.13	.937	.171	
Motivation to reduce energy consumption	direction 1	31	4.16	.638	.115	t(59) = .783, p = .437
	direction 2	30	4.00	.947	.173	

Additionally, the qualitative feedback did not reveal a big difference in the preference between the two versions. The participants in the Direction 1 treatment were commenting less on the need to flip the battery (5/30 participants) and the battery orientation was rather implicit by the comments they have made: *"you reach the goal, you expect it to be on top", "energy should fill up from the bottom, and not from the top".* In the Direction 2 treatment, however, they gave more explicit suggestions as to the need of flipping the battery (7/30 participants): *"flip the battery", "The display seems to be inverted in the current*

design", "I would expect energy depletion to start at the top, not at the bottom", "quite confusing at first because the battery depletes from the bottom".

Finally, some of the variables we have coded in order to test the concept understandability behind the visualization, presented in the section "Evaluating the understandability of the concept behind the battery visualization" that can be found below, revealed several significant differences in the understandability of the concept which is better in the Direction 1 than in the Direction 2 treatment. Taking all of the above-mentioned evidence, we decide to use Direction 1 visualization as the main visualization in the enCOMPASS application.

Elaborating the evaluation of the Direction 1 Visualization

As we have decided to focus on the Direction 1 visualization to use in the encompass application, let us look more closely at how users evaluated it. Again, coming back to the question about the clarity of the visualization in the beginning and the end of the survey, we can see in Figure 5.8 that first of all, most of the respondents in the Direction 1 treatment found the visualization to be quite or very clear. Also, we can see that there is an increase in clarity of the visualization in the end vis-à-vis the beginning of the survey.

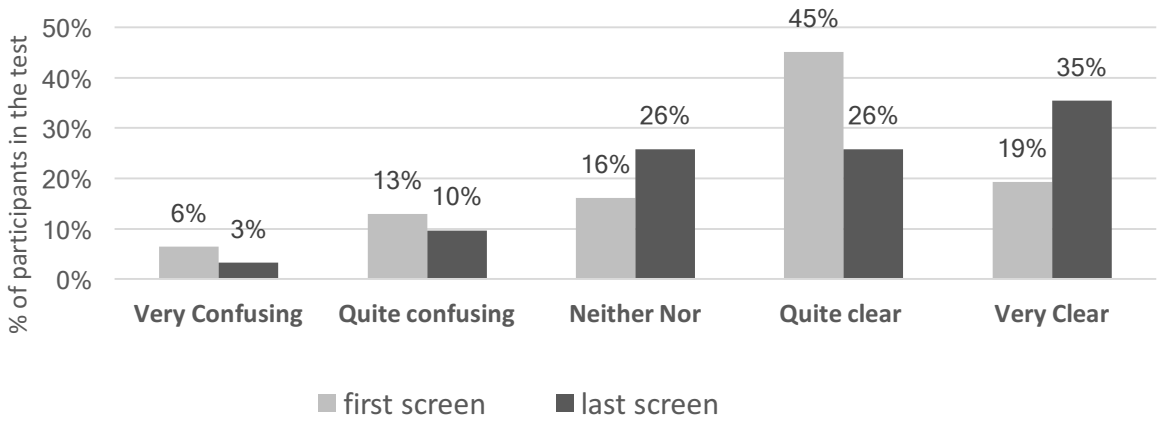


Figure 5.8 Evaluation of confusing-clear dimension of the visualization in the Direction 1 Treatment

Figure 5.9 displays vividly how participants evaluated the Direction 1 visualization regarding the dimensions given by Hassenzahl (2004). Overall, the participants evaluated the visualization positively: 81% gave a score of 7 or higher. The pragmatic quality of the visualization was evaluated very well: 68% of participants gave it a score of 7 or higher, whereas the hedonic quality received somewhat lower notes, but still a majority of participants rating it above average.

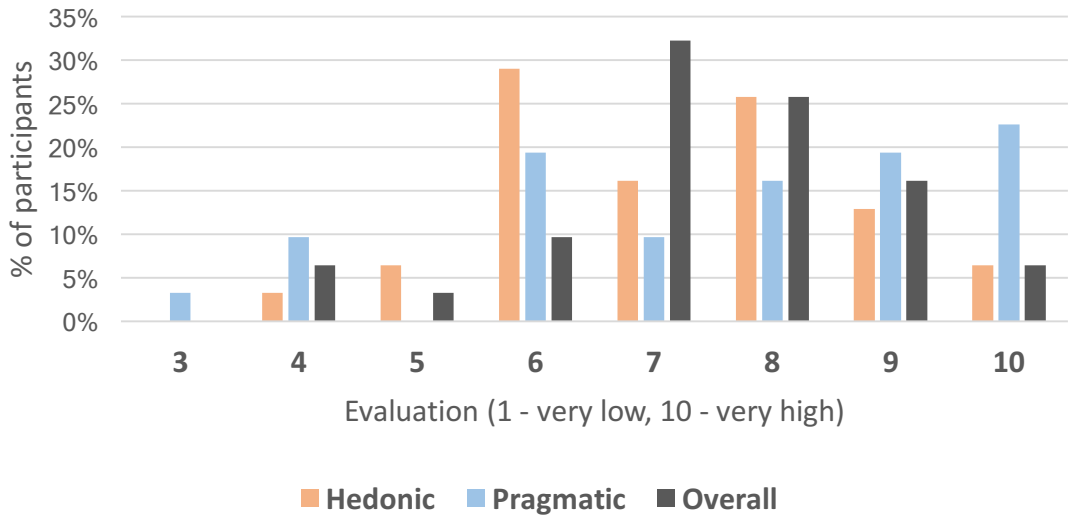


Figure 5.9 Distribution of the evaluation dimensions of Hassenzahl (2004) by participants in the Direction 1 Treatment

Looking at the Figure 5.10 which presents the evaluations of participants according the dimensions of Voss et al. (2003) we can observe a similar result: participants find that the Direction 1 visualization has rather more utilitarian value than hedonic one: 74% evaluated the visualization as being very much or slightly effective vs. only 42% evaluated it to be very much fun or slightly fun.

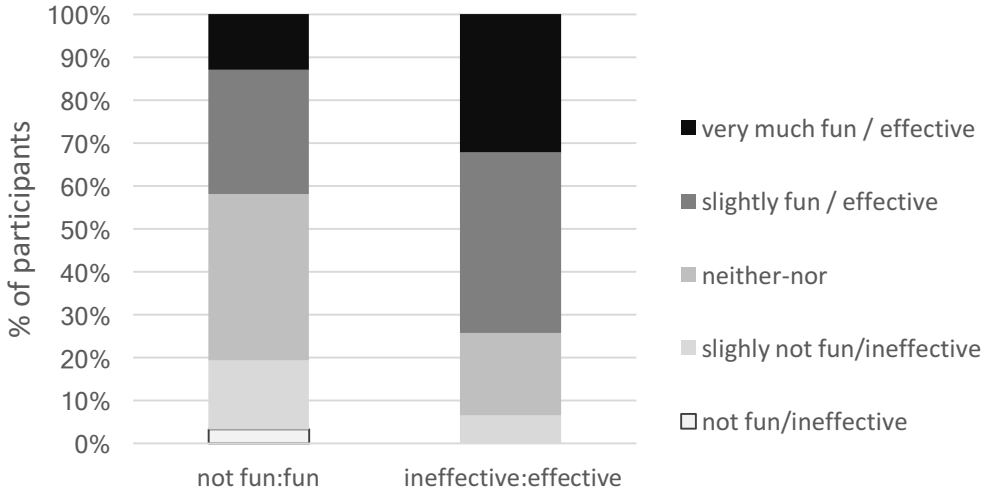


Figure 5.10 Distribution of the evaluation dimensions by Voss et al. (2003) by participants in the Direction 1 Treatment

Finally, in Figure 5.11 we can see how many participants felt motivated by the Direction 1 visualization to meet their energy saving goal or to decrease their energy consumption. We can see that overall 94% of participants agreed strongly or slightly with the statements in questions 19 and 20 in Appendix 2. We therefore conclude that the visualization we have developed allows people to achieve the main goals in the enCOMPASS application – save energy.

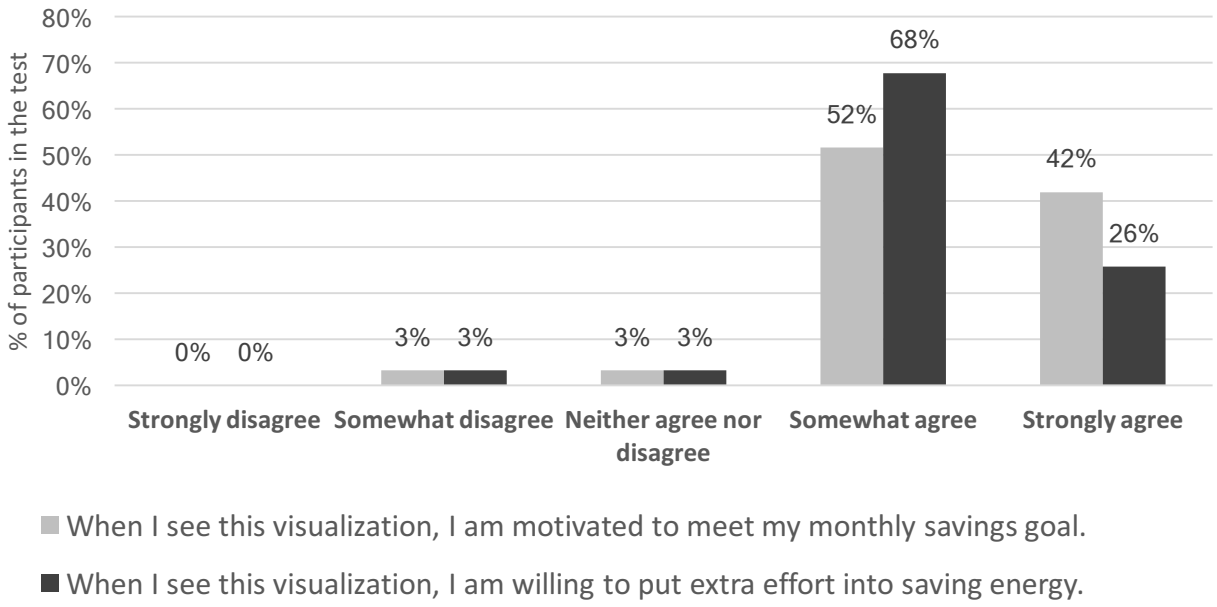


Figure 5.11 Motivation of the user to save energy after looking at the visualization in the Direction 1 Treatment.

Evaluating the understandability of the concept behind the battery visualization

As can be seen in Appendix 2 we have asked users questions about their understandability of the concept behind using the battery metaphor to illustrate their energy consumption and help them to save energy. Although the answers to some of these questions were provided on the open-ended basis, there was usually a right answer implied, at least by the design of the visualization. As a result, their answers were coded into three categories – 1 was given if the answer was completely correct, 0,5 was given if the answer was partially correct and 0 was given if the answer was totally wrong. In the Figure 5.12 below the distribution of the answers is provided for the participants in the Direction 1 treatment (the distributions for the participants in the direction 2 treatment can be found in the Appendix 2). Already judging by this figure, we can conclude that the participants understood most elements of the concept fully or partially. Only very few participants were confused then looking at our visualizations and did not understand the elements of the concept. In the following we will discuss each of these elements in detail and will present the examples of the given right and wrong answers.

The application has some inherent dynamics: the users will be not only able to see their consumption and forecast their ability to meet their savings goal in the current month, but also to see how they performed in the months before. Thus, the questions about understandability of different elements of the visualization relate to both of these aspects. Specifically, the first four questions relate to the visualization of the current month and the progress within that month, whereas the next three questions relate to understanding the result of the previous months and using this information to infer the possibility to reach the goal in the current month, and the last three questions refer to both: the current month and the previous ones because they relate to color coding that is the same with the process vs. the result visualization.

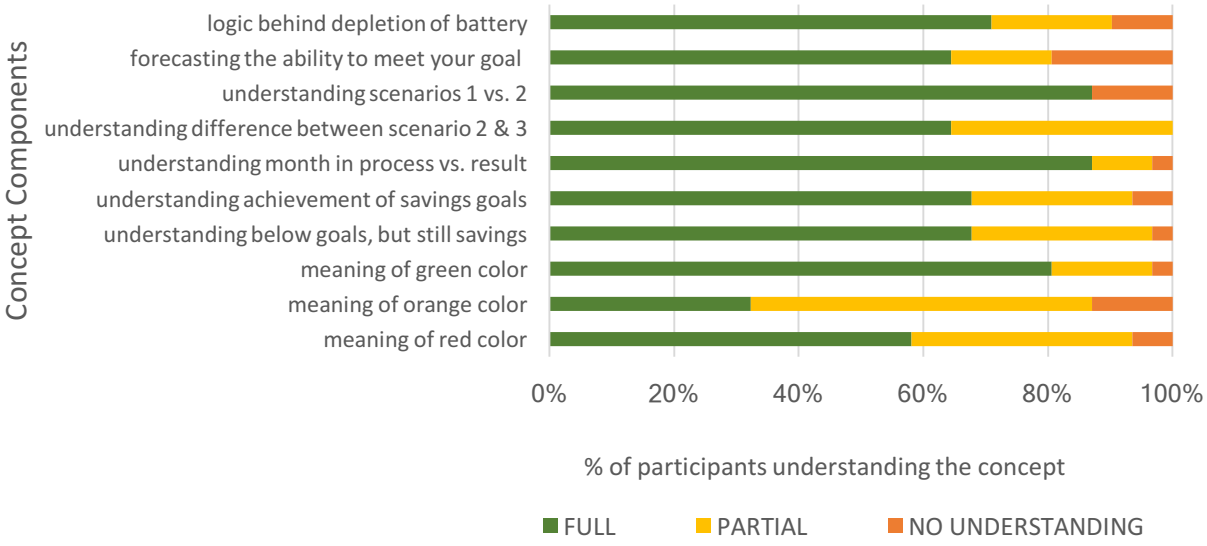


Figure 5.12 Distribution of understandability of different elements in the visualization by participants in the Direction 1 Treatment

First, we wanted to check if participants understand the logic behind depletion of battery when they consume electricity (referring to the first screen of the visualization and question 2 in Appendix 2: What do you think will happen if you continue to use electricity this month? How this would be reflected in the battery visualization?). Full answers were given by 71% of participants and another 19% answered partially correct. An answer by participant similar to: *"The yellow part would get bigger and the green part would get smaller. The monthly use number would increase and the available would get smaller."* Partially scoring answers would look something like this: *"The Battery would gradually be drained out and eventually reach the savings goal. I guess it would eventually turn red if the savings were reached."* This answer is partially correct as the user did not name any elements that would change in the battery and also was speculating on parts that were not asked. And totally wrong answers sounded similar to this: *"I would save at least 5kWh for this month. Although I don't consume more than 75kWh, I find it interesting"* as this answer does not relate to the question at all and mixes up the elements presented in the visualization. The Pearson chi-square $\chi(2) = 6.67$, $p = .0354$ reveals that there is statistically significant association between the understandability of the logic and the treatment group, in the Direction 1 group people understood it better than in the Direction 2 group.

Second, we wanted to understand if the users could forecast their ability to meet their energy savings goal by looking at the visualization once a certain number of weeks have passed in the month (referring to the second screen in the visualization when 2 weeks have passed and to question 7 in Appendix 2: How likely are you to meet your goal if you continue to use the same amount of electricity over the next two weeks?). Full answers were given by 65% of participants and another 16% answered partially correct. People who received full points for answering this question said: *"If I use the same amount I'll reach 90 kw. My goal is to use only 85 kw of electricity, so I can't reach the goal if I use the same amount of electricity of the previous 2 weeks".* Partially correct answers were referring to a possible reduction in energy consumption in the second two weeks, because consumption of week 2 was smaller, however the premise of the question was that the consumption will stay the same: *"I assume I will be successful because I've been gradually reducing*

[4] - We use Pearson Chi-square to evaluate the significance of the differences, because our data is nominal (1 – fully right answer; 0,5 – partially right answer; 0 – wrong answer)

my energy use. If the tendency is repeated I will progressively reduce the usage". Wrong answers mainly happened when people did not refer to the saving goal, but to the rest capacity of the battery: *"We are very likely to meet our goal. In two weeks 45% has been used which leaves us 55% for the next two months (weeks)"*. The Pearson chi-square $\chi(2) = 8.630$, $p = .013$ reveals that there is statistically significant association between the understandability of forecasting the ability to meet the goal and the treatment group, that is in the Direction 1 group people understood this part of the concept better than those in the Direction 2 group.

Third, understanding scenarios 1 vs. 2 aimed at testing whether users can identify when they are still on the right track vs. when they are already past their energy saving goal (refer to the comparison screens of Scenario 1 vs. Scenario 2 and related question 9 in Appendix 2: In which of the two scenarios are you on the right track to meet your goal). This was a closed question, the participants had to choose between suggested options. The right answer - given by 87% of participants - is the Scenario 1, as in Scenario 2 it is vivid in terms of color (orange) and also visually that the user's consumption has surpassed the goal. The Pearson chi-square $\chi(1) = 0.126$, $p = .722$ reveals that there is no statistically significant association between understanding the goal setting in scenario 1 vs. 2; in fact, participants in both treatments understood this concept element equally well.

Fourth, similarly to previous question understanding the difference between scenarios 2&3 aimed at testing whether the user can identify when they are not only past their goal, but also, they are consuming more than in the previous year (referring to the comparison screen between scenarios 2&3 and question 13 in Appendix 2). This question was also testing the understandability of the implicit baseline – that their consumption was compared to consumption in the same month of the previous year. Full answers were given by 65% of participants and another 35% answered partially correct. Fully correct answers would not only recognize that the user is over using the electricity in Scenario 3, but also recognize the baseline to which this relates to: *"In Scenario 3 I have over-used the energy, and consumed more than a year ago. In Scenario 2 I have exceeded my goal, but still stay in limits"*. The participants who answered partially correct, were usually making the mistake thinking that 100kWh is all the energy they have been allocated which is not entirely true: *"In picture 1 the person hasn't achieved their goal however they still have some energy left to consume. In picture two the person hasn't achieved their goal, and they also have no energy left for the month as they have used too much"*. The Pearson chi-square $\chi(1) = 0.208$, $p = .648$ reveals that there is no statistically significant association between understanding differences in scenarios 2 and 3; in fact participants in both treatments understood the differences equally well.

The next series of questions relates to understanding the difference between the current month view and the result of the previous months view. First, in order to differentiate between these two, we have used a different color scheme: to indicate performance during the month we use light green, whereas to indicate the result we use darker green. We wanted to test whether users also see this difference (refer to Question 17 in Appendix 2). Fully correct answers were given by 87% of participants and another 9% answered partially correct. The correct answers would look like: *"Light green is still in progress and bright green means the month has surpassed and the goal has been hit"*. This participant received partial points for understanding the meaning only of the part of the color coding scheme: *"The dark green is well over the savings goal and is safe. The light green is very close to the savings goal and is at risk of spending over"*. The Pearson chi-square $\chi(2) = 3.627$, $p = .163$ reveals that there is no statistically significant association between understanding the difference in the process vs. result; in fact, participants in both treatments understood the difference between process and result equally well.

Second, we wanted to test if users can recognize when they were viewing the results of the past months if they recognize those when they achieve their saving goals (refer to question 15 in Appendix 2). Most of the

users (68%) could recognize that they only achieved their saving goal in August, whereas another 26% named along August also September (or another month) as they did not understand that you can only achieve the goal once the month has passed. The Pearson chi-square $\chi(2) = 0.407$, $p = .816$ reveals that there is no statistically significant association between understanding achievement of savings goals and the treatment group, that is in both groups understandability was the same.

Third, we wanted to test whether the users can recognize the months when they reduced their consumption as compared to the same month in the previous year (question 18 in Appendix 2). Most of the users (68%) could recognize that they managed to reduce consumption in both June and August, whereas another 29% named only one of these months. The Pearson chi-square $\chi(2) = 1.215$, $p = .545$ reveals that there is no statistically significant association between the understanding achievement of savings goals and the treatment group, that is in both groups understandability was the same.

Finally, in this set of questions we wanted to test whether users could understand the meaning of the color scheme. First, they were asked about the meaning of the color green when they were presented with the two different scenarios for the first time (refer to question 11 in Appendix 2). 81% of participants could recognize that the color green means *"You are still above your goal so it will remain green until you go under it"*, whereas another 16% related this rather to the result than the process: *"You reached your goal of consuming 15% less energy than last year"*. The Pearson chi-square $\chi(2) = 1.024$, $p = .599$ reveals that there is no statistically significant association between understanding the meaning of the color green and the treatment; in fact, participants in both treatments understood the meaning of the color green equally well.

Second, they were asked about the meaning of the color orange when they were presented with the two different scenarios for the first time (refer to question 10 in Appendix 2). Only 32% of participants could recognize that the color orange means that *"We are over our target but below last year's usage"*, whereas most of participants (specifically, 55%) correctly recognized that they are past the goal, but did not mention that they are still above the last year's usage. The Pearson chi-square $\chi(2) = 5.010$, $p = .082$ reveals that there is no statistically significant association between understanding the meaning of the color orange and the treatment; in fact, participants in both treatments understood the meaning of the color orange equally.

Third, they were asked about the meaning of the color red when they were presented with scenarios 2 and 3 (refer to question 14 in Appendix 2). 58% of participants could recognize that the color red means *"It means that I have missed my goal and used more electricity than the previous year"*, whereas another 35% made a small mistake in their judgment e.g. relating to not understanding the baseline: *"no energy left to use that month"*. The Pearson chi-square $\chi(2) = 1.567$, $p = .457$ reveals that there is no statistically significant association between understanding the meaning of the color red and the treatment; in fact participants in both treatments understood the meaning of the color red equally well.

During the survey, the participants were also answering a series of True/False statements (see questions 1, 6 and 14 in Appendix 2) in order to double-check their understandability and focus on some specific elements that they might not have mentioned in their open-ended answers. The distribution of their answers to these statements is presented in Figure 5.13. We can see that most of the participants answered most questions as we expected, except for the questions 1.2 and 14.1. We later noticed that there is a small problem with the formulation of these questions. The statement 1.2 reads: *This month I cannot use more than 100kWh of electricity*. This statement was supposed to measure the understandability of the baseline measurement in the battery: the battery displayed in the visualization had a capacity of 100kWh because this was the consumption of the previous year. However, users were able to spend more than that amount. Just changing the wording to "should" would actually make the statement true, because users were gauged in the visualization to spend not as much energy as their baseline. As we

do not know how they interpreted this question, we cannot use their evaluations. The same applies to statement 14.1.

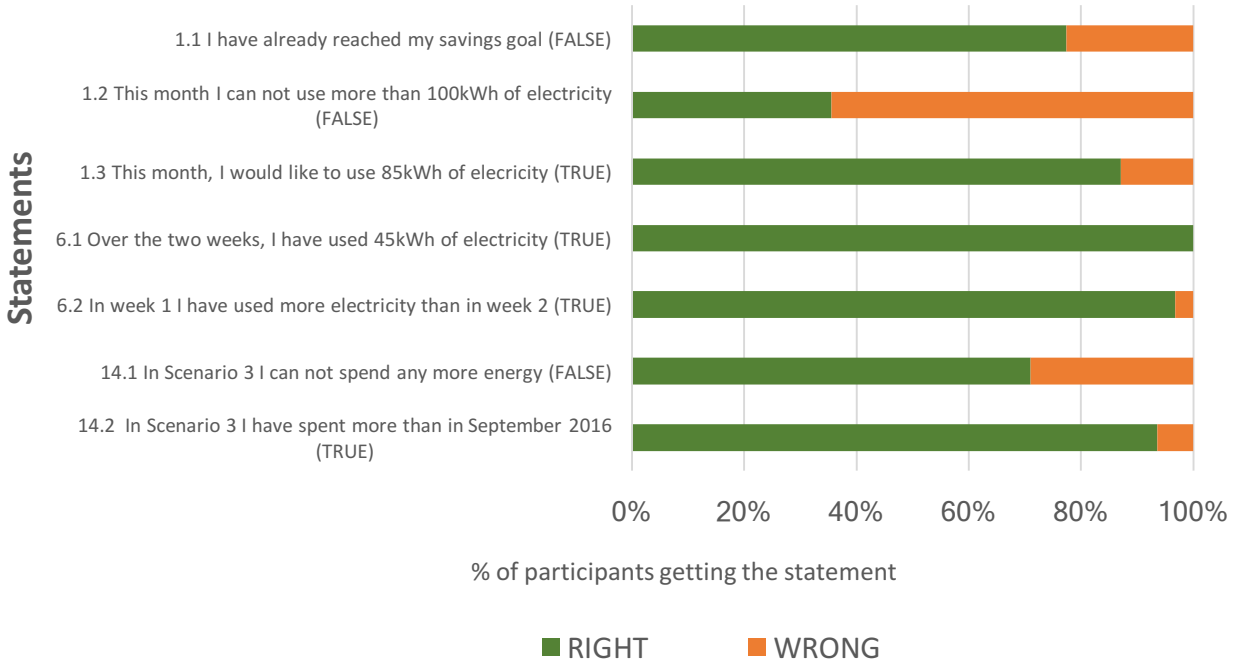


Figure 5.13 Distribution of answers to the True/False statements by participants in the Direction 1 Treatment

There were also no significant differences between the treatments in answering these questions. A series of Pearson chi-square tests reveals that there is no statistically significant association between answering statements correctly and the treatment; in fact, participants in both treatments had similar probability to answer the statement correctly; this is valid for statements 1.1, 1.3, and 14.2 with the respective results $\chi(1) = 3.070$, $p = .080$; $\chi(1) = 1.122$, $p = .289$; and $\chi(1) = 0.317$, $p = .648$[5].

Understandability of selected UI elements

Some questions in the survey presented in Appendix 2 related to specific UI elements in the visualization. This was the first test to determine if the users understand what clicking on certain buttons would do or state change of buttons would signify. This was the most challenging part for participants, as they were presented with the static images and could not try out these elements in the application. The results presented in the Figure 5.14 reveal that some elements were harder for participants to understand than others, and considering the fact that it is hard to evaluate usability of UI elements without actually trying them out, these results should be taken with a grain of salt. However, these insights were valuable to adjust some elements of our designs which were not easily understood by users.

First, referring to the question 12 in Appendix 2 the “edit goal” button would change from green to grey because the user could not change it once the energy consumption has exceeded the goal. This was fully recognized by 81% of participants and another 16% of participants gave a good answer which shows

[5] - Question 6.1 is not evaluated as all participants in both treatments answered this question correctly, whereas Question 6.2 is not evaluated because in the direction 1 visualization week 1 consumption was higher than week 2, whereas in the direction 2 visualization it was vice versa.

understandability, but does not actually answer the question fully, such as *"It could be a further color-coded indicator that you have missed your target"*. The Pearson chi-square $\chi(2) = 5.488$, $p = .064$ reveals that there is no statistically significant association between button change grey and the treatment; in fact, participants in both treatments understood the meaning of this UI element equally well.

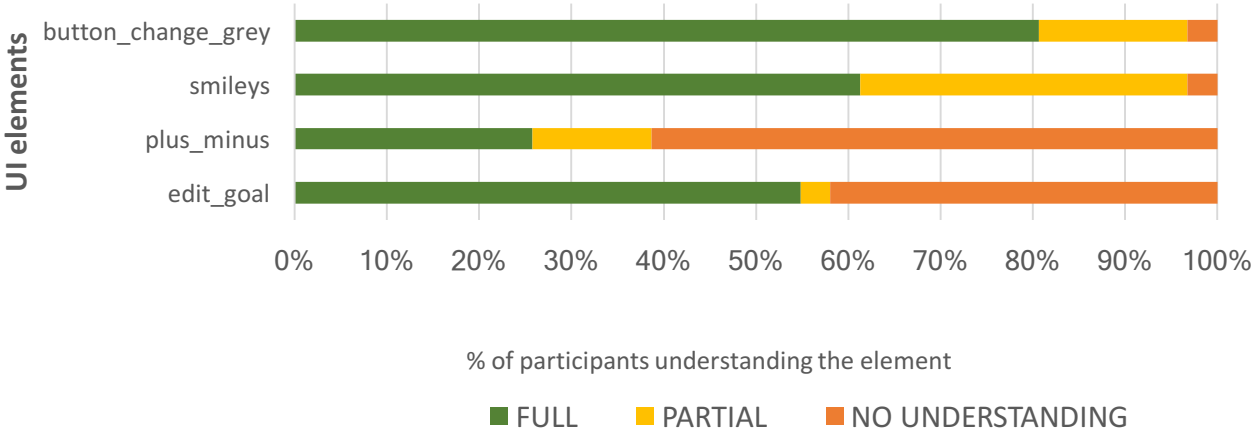


Figure 5.14 Understandability of UI elements by participants in the Direction 1 treatment.

Second, referring to the question 16 in Appendix 2, smileys were given to participants to provide additional indication to the colors about how well a user has performed regarding goal achievement. As 61% of users rightfully recognized: *"The :) means that I made the goal for that month. The :| means that I didn't make the goal but I also didn't use more than the previous year. The :(means that I used more than the previous year"*. Additionally, partial answers (35% of participants) were given if the users did not explain the meaning of all the smileys or broadly referred to the concept: *"They indicate how the app feels about my energy use"*. The Pearson chi-square $\chi(2) = 3.567$, $p = .168$ reveals that there is no statistically significant association between the understandability of smileys and the treatment; in fact, participants in both treatments understood the meaning of this UI element equally well.

A lot of confusion arose due to the meaning of plus and minus signs in the battery (refer to Question 8 in Appendix 2). As 26% of users rightfully pointed out: *"It can be considered just an emulation of a real battery"*. Another 12% of users were not sure whether they meant something or they did not: *"On one side they reflect the poles. On the other hand, they show if I used more or less energy"*. Most of the users, however, tried to find the meaning in the plus and minus signs: *"High amount of electricity available, and low amounts of electricity available"*, but this can also be due to the fact that participants were induced to infer meaning by the framing of the question. The Pearson chi-square $\chi(2) = 6.568$, $p = .037$ reveals that there statistically significant association between the understandability of plus-minus signs and the treatment; in fact more participants in the direction 2 treatment could recognize that these items do not mean much. However, this is not such an important element to be concerned that it compromises the choice of the direction 1 visualization. As most users were confused as to the meaning of these signs, we decided to remove them from the visualization.

Some confusion also arouse around the functionality of the "edit goal" button (refer to Question 3 in Appendix 2). 54% of the users understood the functionality of this button, whereas 42% thought that by clicking on the button they would be able to put in notes possibly due to the fact that the button used the pen icon often used in the writing applications. The Pearson chi-square $\chi(2) = 1.585$, $p = .453$ reveals that there is no statistically significant association between the understandability of the edit goal button and the treatment; in fact, participants in both treatments understood the meaning of this UI element equally well.

As a result, we decided to rethink which icon to use for this button and for the working version of the visualization to use "edit goal" written out.

Qualitative feedback provided by participants

In the end of the survey, the users were asked two open-ended questions about what users liked and disliked about the visualizations they were presented with (refer to questions 22 and 23 in Appendix 2). We have coded their answers and calculated the frequencies. As the answers were open-ended and the users were not prompted to answer in a certain way, the diversity of their answers is quite high and therefore, the frequencies are quite low. However, we will mention the most frequent/interesting issues they have raised.

On the positive side, a lot of people said that they found the visualization is simple, clear and easy to understand (32% in the direction 1 30% in the direction 2 treatment): *"I like that it is simple, easy to understand and I think it could be helpful to many people. Even a kid would understand every concept, which is a good thing for this kind of market".* The users also described the visualization as: appealing, adjustable, motivating and informative. They liked using the concept of the battery to display energy savings (10% in the direction 1 and 20% in the direction 2 treatment): *"I like that the visualisation feels quite familiar having been used to the battery usage icons on mobile cell phones".* Many noted a minimal and simple design (26% in the direction 1 and 23% in the direction 2 treatment). The elements that the users liked the most were the color coding scheme (45% in the direction 1 and 57% in the direction 2 treatment) and the smiley indicators (19% in the direction 1 and 17% in the direction 2 treatment): *"I also like the traffic cone and smiley indicators as it is clear and easy to immediately identify how your usage is going for the month is going in relation to the goals set".* They also mentioned such features as the weekly progress, comparison to last year, the ability to go back in the other months of the year. Overall users said that the visualization would allow them to achieve their main purpose - adjust their energy savings: *"You have plenty of warning if your energy use is too high in the first few weeks in order to change things before the target is exceeded".*

On the improvement side, people mentioned that they had to struggle with understanding the visualization initially (13% in the direction 1 and 20% in the direction 2 treatment): *"The visualization took some time to understand. Firstly, it was not clear, what the purpose of it is".* Some users mentioned the design being too simple and lacking animation or features (19% in the direction 1 and 10% in the direction 2 treatment) and made some suggestions for its improvement, such as: use more white space, increase the font at some places, change the yellow color of electricity used, use different colors to distinguish process and result, display the figures on the battery itself and not in the boxes next to it. All of these issues are minor and can be improved in the final design stage. Also participants saw improvements in the concept by adding extra features such as: setting goals in between, displaying daily use of electricity, showing consumption for all of the months in a year on the same view, showing over-consumption in %, setting goals also in Kwh along the %, displaying energy usage trends. Some users had some confusion about the baseline being the last year and suggested using the goal as a baseline: *"Some people may be silly enough to think that if they use all the energy compared to last year then they will not be able to use any more."* Of course, these are nice features, however the visualization is yet complex enough to add more layers to it or to think it over again. However, smaller things, such as showing goals in KwH or over consumption in % can be implemented.

Some people said that they disliked nothing in the visualization (16% in the direction 1 and 10% in the direction 2 treatment). Few people (10% in the direction 1, none in the direction 2) wished for gamification elements, which will in fact be present in the application: "*Also, even though as I said I think it's quite challenging, it's not too "funny" and I would love to see some sort of small achievements".* Others were missing the information on how to save more (13% in the direction 1 and 10% in the direction 2): *"It doesn't*

say how exactly you can and do save energy and how this is measured" – again referring to an element which was missing in the test but will be present in the enCompass application.

5.2 RESULTS FROM IMPACT VISUALIZATIONS TESTS

5.2.1 Methodology

In order to test the understandability of the impact visualizations presented in Section 3.3, we conducted a survey among possible users of the application. The survey was developed in an iterative process and aimed to test several things: 1) the conceptual understandability of the impact visualizations; 2) the evaluation of the visualizations along the usability dimensions, and 3) the preference towards a certain type of an impact visualization. Similar to the overview visualization test, the users were presented with the screenshots of the visualization and asked a series of questions. Sometimes it was difficult to work with static images as the application usage assumes certain dynamics. To illustrate dynamics, we have utilized animated gifs and provided instructions to the users to imagine themselves using this application in a specific context.

The participants were recruited through Amazon Mechanical Turk as with the overview visualizations. For all of the tests we have conducted, similar instructions as presented in Appendix 3 were given to participants in these tests. The qualification criteria were also the same (from Europe, approval rate higher than 95% and it was especially important that the users did not take part in our other survey). The survey was conducted through Qualtrics. The format of the questions was similar to the one used for the Overview visualizations. See section 5.1.1 for details. We have used same questions where possible, and also used same items to measure the attitude towards the visualizations. We also added questions in the beginning of the test to measure respondent's overall values (environmental, hedonic and egoistic[6]) with items by Steg et al. (2013).

Following the common methodology for crowd-tests, we have first conducted several pre-tests with a small group of 10 participants to identify upfront any understandability issues for the visualizations and the questionnaire, and any obvious improvement potential. Based on the feedback from the pre-tests we adapted the interim visualizations and the questionnaire accordingly and then performed the actual tests with a larger group of participants. The final version of the survey can be found in Appendix 6.

Testing all three visualizations at once was unfeasible due to the duration the test, which carried the risk of participants dropping out without completing it. Similarly, time to recruit enough participants and have them complete the test was an issue, as was synchronizing the tests with development timeline of the individual visualizations. Accordingly, we performed test of the environmental and monetary impact visualization in one test, as soon as they were available (test 1), followed by the subsequent test of the hedonic visualization (test 2). The combined test of the environmental and monetary impact visualization – test 1 – included 38 questions (refer to Appendix 6). First participants answered questions about the monetary impact visualization, and then about the environmental visualization. The participants were acquired in two batches: in the first batch 11 people filled it out and took on average 19 minutes, in the second batch 22 people filled it out and took on average 23 minutes. In total 33 people took part in test 1 and in both batches the users were rewarded with $1.25 for their efforts. The test of the hedonic visualization – test 2 – included 20 questions (refer to questions marked with * in Appendix 6) and was

[6] - The original item list by Steg et al. (2013) also includes altruistic value orientation, but we did not measure this one as it does not relate to our concept.

completed by 32 participants, who took on average 14 minutes to complete the test. They were rewarded with $0.75 for their efforts (since this was less effort than doing two tests in a row).

In the following, we present the demographics of the users who took part in test 1 and test 2. The demographics are quite similar to the demographics of the participants in the overview visualization tests presented in Section 5.1.2: majority of the users is young (see Figure 5.15), but we also see users from other age groups, there is more balance what concerns gender distribution (Figure 5.16) than in the overview visualization test, the respondents are quite educated (Figure 5.17) and have their household sizes are evenly distributed (Figure 5.18).

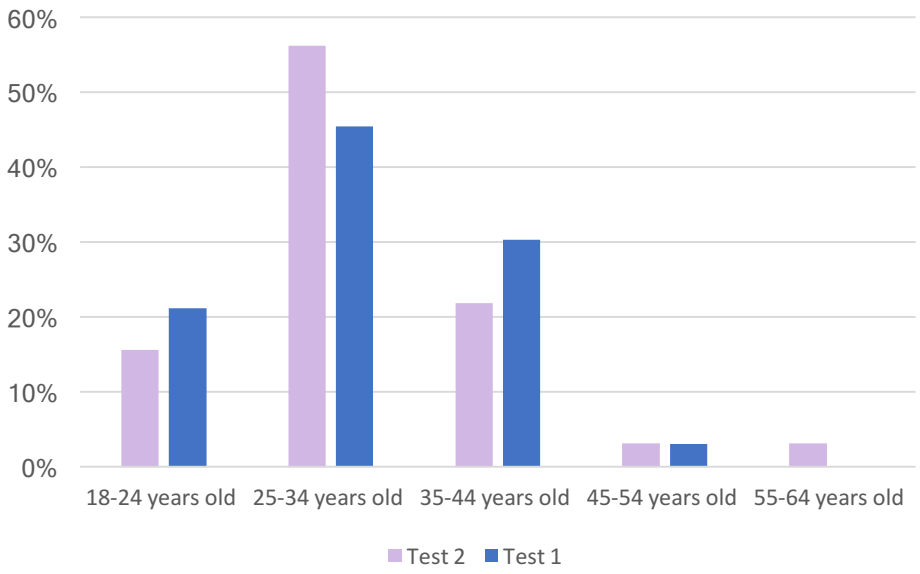


Figure 5.15 Country of Origin of respondents in the test

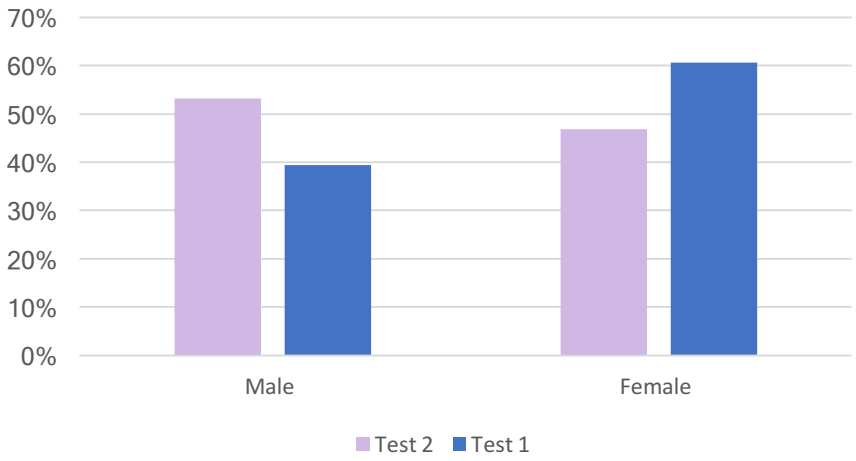


Figure 5.16 Gender of respondents in the test

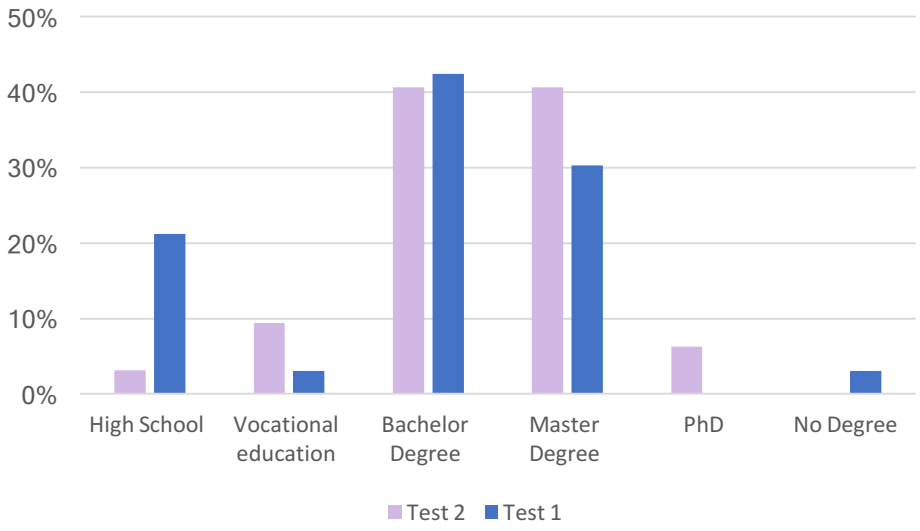


Figure 5.17 Educational Level of respondents.

Figure 5.18 Number of people in the household.

Additionally, in Figure 5.19 we present values reported by respondents according to the classification provided by Steg et al. (2013). In the figure, we can see similar dynamics with regards to values between the participants of test 1 and test 2: a bit over 40% of people rate environmental values as very important or higher, almost 80% rate hedonic values as very important or higher and only 10% in test 2 and almost 20% in test 1 rate egoistic values as very important or higher. Egoistic values include items which directly measure directly user's interest in material possessions and money. This distribution of values supports our theory-grounded choice of the three different types of impact visualizations (discussed in Section 3.1).

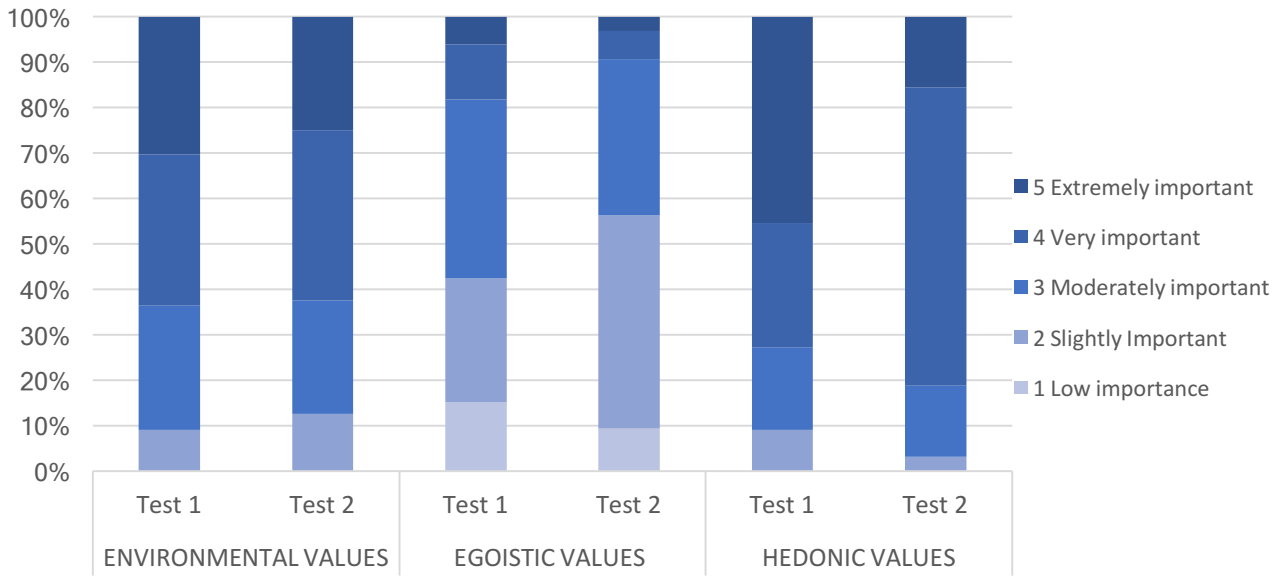


Figure 5.19 Environmental, Egoistic and Hedonic values of the respondents as measured by items of Steg et al. (2013)

5.2.2 Test results and user feedback

Main aims of the overview visualization test were (in order of importance): 1) the conceptual understandability of the impact visualizations; 2) the evaluation of the visualization along the usability dimensions, and 3) the preference towards a certain type of an impact visualization. We will thus present the results in this order as well. In the following we present the results for the monetary and environmental visualizations which were carried out in test 1 as well as the hedonic visualization which were obtained through test 2 discussed in 5.2.1. Please note that in this section we will not be presenting any statistically significant results, as our aim is not to compare between the groups, but just look at the different dynamics.

Concept Understandability of impact visualizations

As can be seen in Appendix 6 we have asked users questions about the understandability of the metaphors that we have used to display the impact of their energy savings. Although the answers to some of these questions were provided on the open-ended basis, there was usually a right answer implied, at least by the design of the visualization. As a result, their answers were coded into three categories: – 1 was given if the answer was completely correct, 0,5 was given if the answer was partially correct and 0 was given if the answer was totally wrong. In the Figure 5.20 the distribution of the answers to these questions is provided when participants were presented with environmental and monetary visualizations (same participants) and hedonic visualization (a different group of participants as discussed in 5.2.1). Already judging by this figure, we can conclude that the participants understood most elements of the impact visualizations fully or partially. In the following we will discuss each of these elements in detail and will present the examples of the given answers.

First, we wanted to check if participants focused on the main graphical elements that they were presented with (referring to the first screen of the respective visualization and question 2 in Appendix 6: Please name the graphical elements in this screenshot). Full and partially full answers were given by 73%, 73% and 56% of participants about the monetary, environmental and hedonic impact visualizations, respectively. Another

21%, 27% and 41% gave a partially correct answer, which is especially notable with the last hedonic visualization. Here instead of the expected "ball, jar and rewards" the participants would give the same answer twice, such as naming both bubbles and particles, and missing the rewards. It can be expected, as rewards badges were quite small compared to other images.

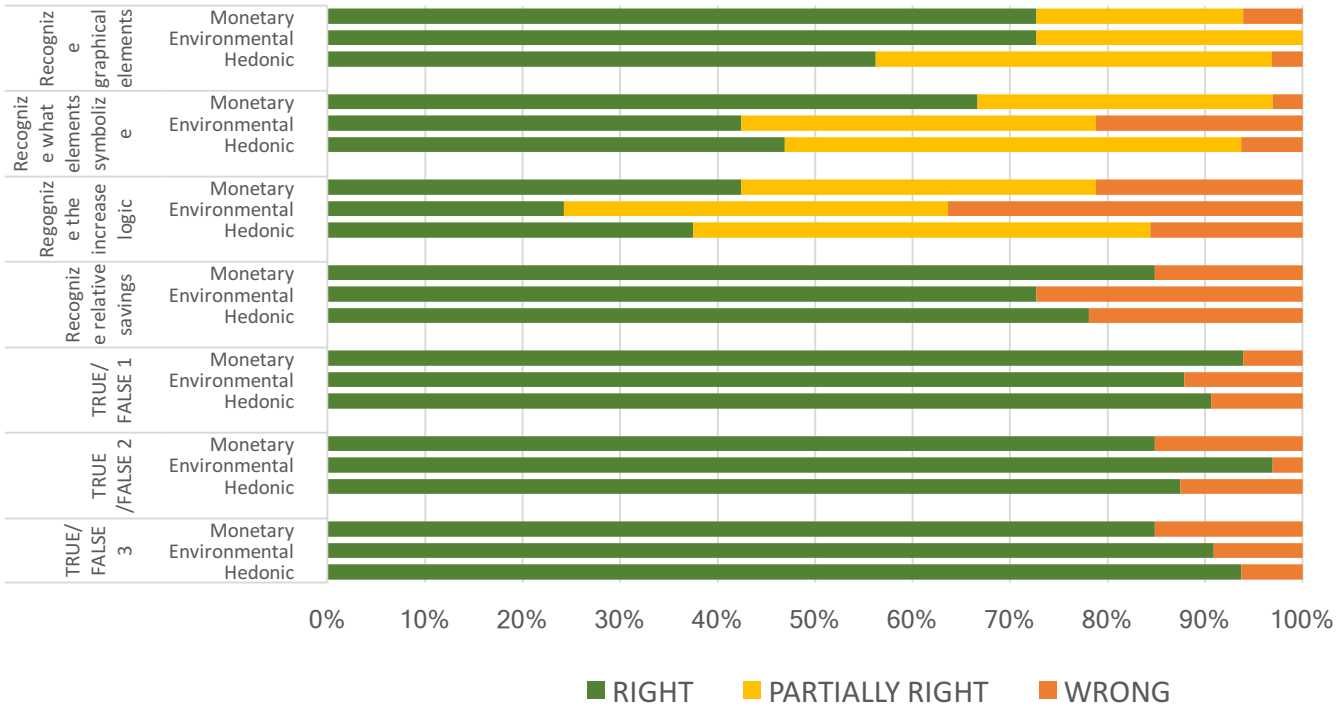


Figure 5.20 Concept understandability of the Monetary, Environmental and Hedonic impact visualizations

The second question dealt with a deeper understanding of the meaning behind the graphical elements and asked the respondents what they thought the graphical elements represented (referring to the question 3 in Appendix 6). Here 67%, 42% and 47% of participants on the monetary, environmental and hedonic impact visualizations, respectively gave full answers and another 30%, 36% and 47% gave partially correct answers. The partially correct answers usually did not recognize what the images were trying to convey, e.g. environmental or monetary savings directly, but focused on some very specific part of the image. For example, one participant explained the following about the environmental visualization – just explaining what she was seeing and not what the visualization meant: *"The cloud represents 1 Kg of CO2 that was not released into the air thanks to a lower consumption. A tree represents 22 kg of CO2 not released in the air"*.

The next question dealt with the ability of participants to recognize the dynamics inherent in the cumulative accumulation of savings over time (referring to the question 4 in Appendix 6).

). As we can see in the Figure 5.20 above 42%, 24% and 38% of participants could answer the question completely right and another 36%, 39% and 47% partially right about the monetary, environmental and hedonic visualizations, respectively. Although most could easily recognize the idea about the temporal change between the screens, in all of the visualizations participants struggled to understand that the savings presented were not about the month in question, but cumulative: *"As I click on the arrow I can select single months. The tree picture adjusts to show how much pollution I avoided in the selected month"*. This information although presented on the visualization, might have been overlooked by the users or possibly they just expressed their thoughts in this way, but actually meant cumulative savings. Also, it might

be the case that if participants had the possibility to experience the application and it was connected to their actual energy use and savings, they might have been able to recognize this easier.

The next question relates to assessing the relative amount of savings that the users have achieved (referring to the question 7 in Appendix 6). Most of the participants answered this question correctly: 85%, 73% and 78% of participants evaluating monetary, environmental and hedonic impact visualizations, respectively. Some participants, however, mentioned that they were missing the information about their overall potential savings.

The last three questions in the Figure 5.20 refer to the TRUE/FALSE statements aiming to understand the logic behind the accumulation of savings and determining relative amounts of savings added by the different months (referring to the question 5 in Appendix 6). Most of the participants answered these questions correctly regarding all three visualizations, which reflects good understandability of the concept.

As the hedonic visualization contained another layer, four more questions assessed participants' understandability regarding this type of the impact visualization. The results presented in Figure 5.21 reveal that most participants answered these questions right or partially right, except for the last question. We wanted to check if participants understood the connection between energy savings and the expert saver symbols (referring to the question 4a in Appendix 6) and over 80% of participants understood it right and another 20% partially right. Another question dealt with differentiating between the different levels of savers by means of the different colors (referring to the question 4a in Appendix 6), where again almost 80% of participants answered the question correctly. The last two questions dealt with recognizing the exact saver status they have achieved by a certain time period. Although most participants (81%) could correctly recognize themselves as beginners after several months of using the application (referring to the question 4b in Appendix 6), only 34% could recognize that they were not yet advanced savers after one year, although they started to collect the points of this color (referring to the question 7a in Appendix 6). This is probably due to the small size of the badge and unnoticeable small tick once the status has been reached – we have taken these minor corrections into the design considerations of the final visualization screens.

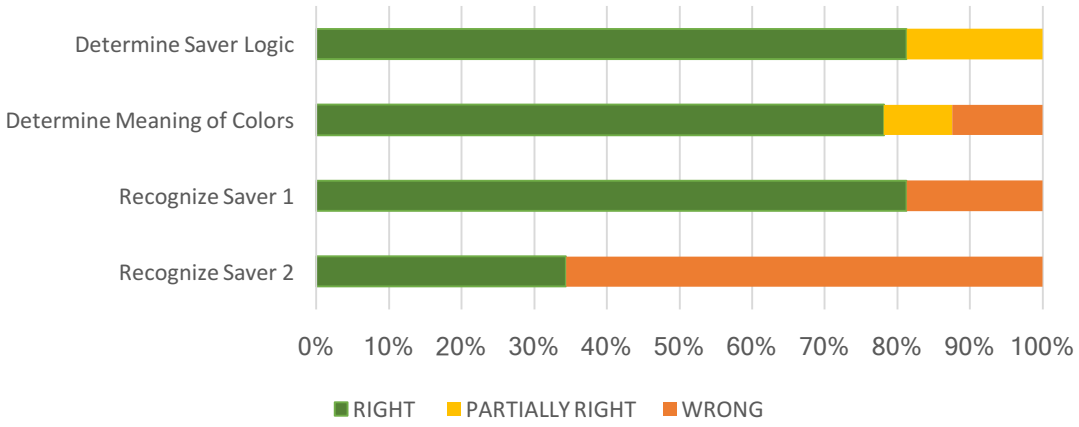


Figure 5.21 Additional items of concept understandability of the Hedonic Visualization

Evaluations of impact visualizations

To evaluate the visualizations and to be able to descriptively compare between them, similar to the overview visualization, we used two approaches, one suggested by Hassenzahl (2004) that identifies two dimensions: the hedonic quality (measured by items such as Clear-confusing and Lame-exciting) and the

pragmatic quality of interactive products (measured by challenging-easy and complicated-simple), complemented by the measurement of beauty (measured by beautiful-ugly). Another approach is the standard marketing approach (Voss et al. 2003) to identify the hedonic (by the item fun – not fun) and utilitarian (effective – ineffective) aspects of consumer's attitudes to products. The questions were asked on a 5pt scale in form of a semantic differential (refer to question 9 in Appendix 6). The average results given by participants regarding the monetary, environmental and hedonic visualization can be seen in the Figure 5.22. We observe that the evaluations are quite similar for most of the dimensions and lie in the range between 2.9 and 3.7. We can see that the monetary and hedonic visualizations are indeed very similarly perceived to be easier to understand, clearer, simpler and more effective than the environmental visualization. However, the environmental visualization was perceived to be slightly more beautiful than the other two. All of the visualizations were perceived to be equally fun and equally exciting.

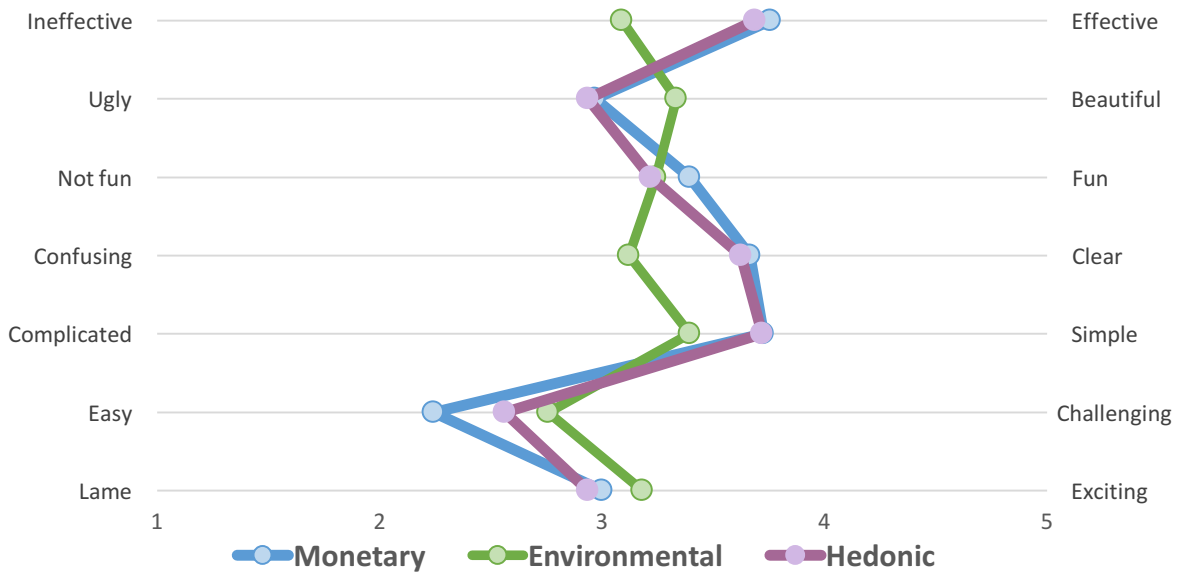


Figure 5.22 Semantic differential evaluation items of the monetary, environmental and hedonic visualization

Figure 5.23 displays how participants evaluated the monetary, environmental and hedonic visualizations regarding the dimensions given by Hassenzahl (2004). We observe some interesting results: participants evaluated the monetary visualization as possessing little hedonic quality (15% of participants gave scores of 4/5 or higher), whereas it possesses a much larger pragmatic quality (over 70% gave 4/5 or higher). Environmental visualization and hedonic visualization have been evaluated similarly: 36% for environmental (and 31% for hedonic) gave scores of 4/5 or higher regarding its hedonic quality; whereas 48% (environmental) and 60% (hedonic) respondents evaluated pragmatic quality with 4/5. However, what concerns beauty, most of participants preferred the environmental visualization (51% gave a score of 4/5 or higher), vs. only 30% gave this score for the monetary and 38% for the hedonic visualization.

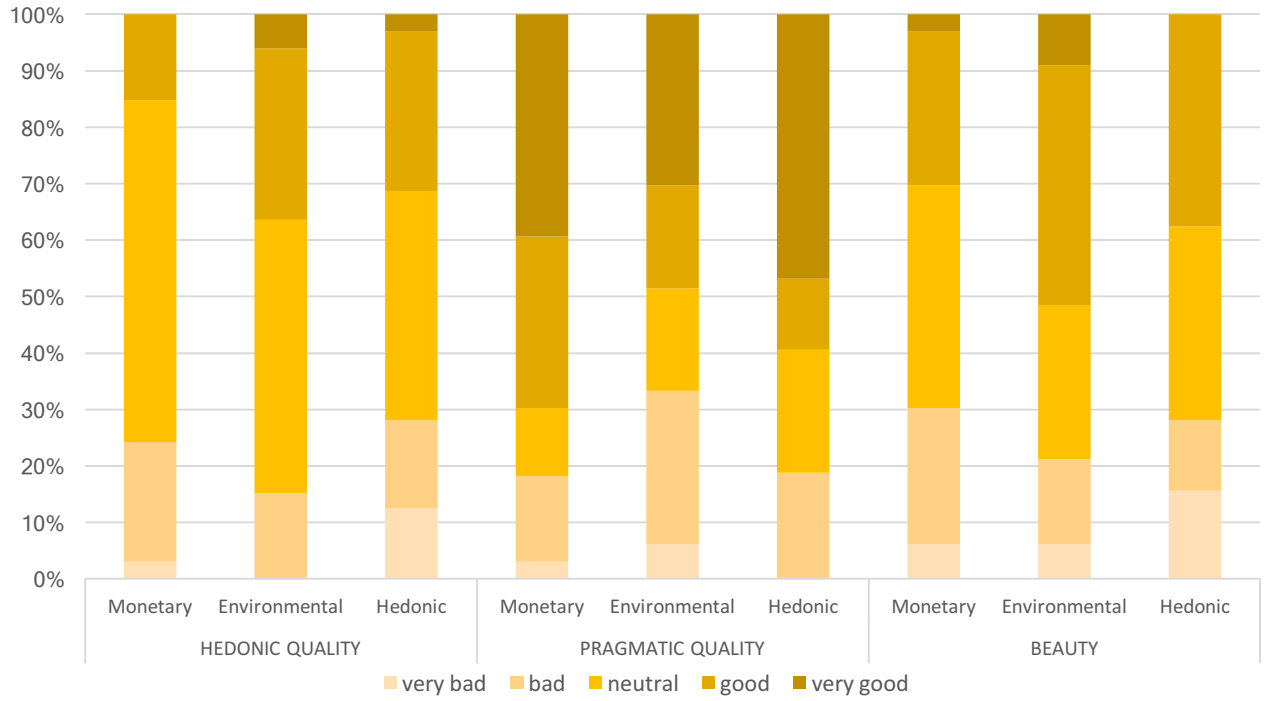


Figure 5.23 Evaluations along the dimensions of Hassenzahl (2004)

Looking at the Figure 5.24 which presents the evaluations of participants according the dimensions of Voss et al. (2003) we can observe a slightly different result as with the dimensions of Hassenzahl (2004) presented in the last paragraph. Specifically, 51% of participants evaluated monetary, 48% environmental and 41% hedonic visualizations with a score of 4/5 or higher on the hedonic attitude dimension, i.e. very similar attitudes. On the other hand, what concerns the utilitarian attitude, hedonic and monetary visualizations were perceived similarly (64% for environmental and 69% for the hedonic visualization have a score of 4/5 or higher), whereas environmental visualization was perceived as slightly less utilitarian (42% gave a score of 4/5 or higher), which could be expected as environmental visualization is trying to promote a less utilitarian, but rather pro-social attitude.

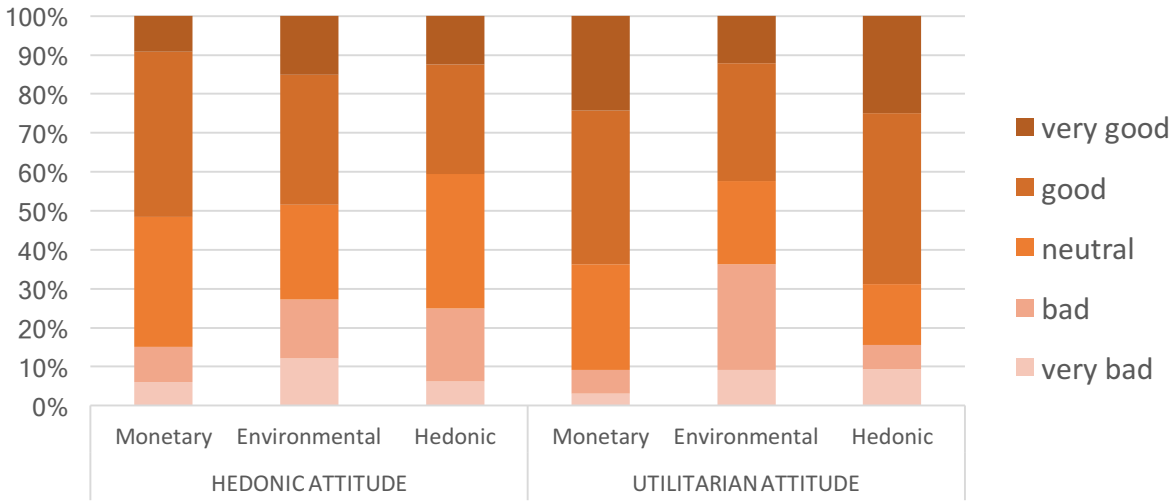


Figure 5.24 Evaluations along the Attitude dimensions of Voss et al. (2003)

Finally, in Figure 5.25 we can see how many participants felt motivated by the impact visualizations to save energy. We can see that the impact visualizations motivate users quite highly and there are no large differences between the motivational power of the visualizations: 82%, 88% and 84% of participants viewing the monetary, environmental and hedonic visualizations, respectively strongly or slightly agreed that the visualizations motivate them to save energy (refer to question 8 in Appendix 6). We therefore conclude that the visualizations we have developed allows people to achieve the main goals in the enCompass application – save energy.

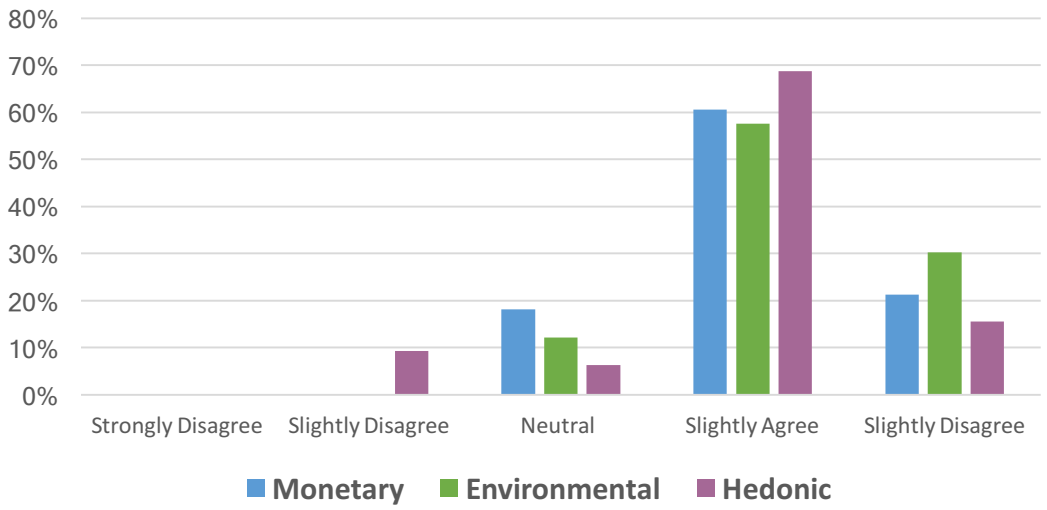


Figure 5.25 Motivation to save energy once the visualization has been seen

In order to shed more light on the aspects about the impact visualizations that the participants liked as well as those they would want to improve, we analyze the qualitative feedback. For each impact visualization, the participants were asked two open-ended questions (refer to questions 10 and 12 in

). We have coded their answers to these questions and calculated the frequencies. The elements that the participants liked about the visualizations were quite uniform so that we can present the results in Figure 5.26. The elements that they would want to improve were quite diverse, so they could not be summarized in a similar figure, but we will discuss in the text the most frequent/interesting issues they have raised.

In Figure 5.26 we can clearly see that the participants found the visualizations to be clear and simple, easy to understand, they have complemented on the design of the elements as well as on making the savings tangible and vivid. We can also observe some differences depending on the visualization that the participants were viewing. The environmental visualization is the leader what concerns the concept – a lot of participants commented on the applicability of the concept: "*I like the idea of having a tree to symbolize my energy savings, as it represents more closely the idea of helping the environment through my actions. It is cool to know how much CO2 we have saved."* The monetary visualization was regarded as very clear and playful and appealing, as well as challenging and engaging: *"I like the way the piggy banks appear, like a challenge to complete it".* The fun visualization scored the most points on the ability to concisely present a lot of information in an engaging way: "*Interactive and very visual. Look easy to use and easy to understand and isn't boring in terms of the saved energy".*

On the improvement side, people mentioned that they needed some time to understand the visualizations, as well as to digest the incoming information: *"It takes some time to figure out how to use it".* They also

found the highlighted element to be a bit misleading as well as they were missing a reference value to help them assess their performance: *"I don't understand why one of the piggy banks is larger than the others and there is no information about the potential amount that can be saved in a giving period".* Some suggested the improvements on the design, such as a different choice of colors, or making some elements bigger as they could be hard to read. In the hedonic visualization, the participants needed more context to make a connection between the earned points and saved energy. We will take these suggestions into account when we will be improving the visualizations to include in the final version of the application.

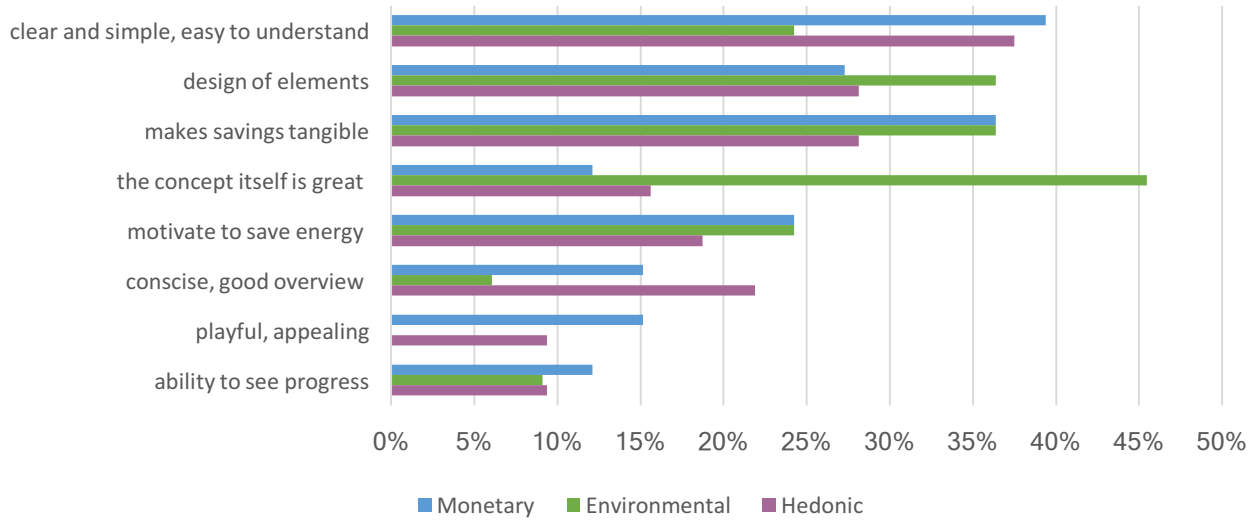


Figure 5.26 Elements of the visualizations that participants liked (qualitative analysis)

5.3 RESULTS FROM AWARENESS APP TEST

This sub section outlines the results of the first preliminary test of integrated awareness app, in which the consumption visualizations developed and described in this deliverable were included. The detailed bar chart view is fully working, whereas snapshots were embedded for the battery overview and the impact views, as alignment with technical partners is ongoing.

5.3.1 Methodology

The awareness app alpha test sought to preliminarily assess the first impression of the participants of the behavioral change incentives embedded in the app. This includes the visualizations, the game features (e.g. leaderboard), and the tips.

Participants were recruited in all pilot locations. Participants were employees and their family from SHF, WVT, and SUPSI who did not have a technical background, are not involved with the project, and have direction 1 technical skills to manipulate a smartphone app. In total 14 participants were recruited (four participants from Germany, three from Greece, and seven from Switzerland).

The test was conducted remotely. Instructions were delivered by e-mail. The instructions contained the tasks the users had to do, which guided them through the application, in order to make sure the users have seen all features. Users were asked to explore the app on their smartphone.

To collect the user's perceptions, a questionnaire was administered, which they were requested to fill out on a desktop computer. The questions contained the following types of questions:

- Application-level technology acceptance questions, derived from Venkatesh et al. (2012), which focused on effort expectancy (i.e. ease-of-use), and performance expectancy (i.e. perceived usefulness).
- Comprehension questions for the visualizations
- Perceived impact of the features on energy saving motivation
- Features of the applications the participants liked, features that could be improved, and improvements that could be made to increase the chance of the participant using the app
- Likelihood of the participants using the application, once it is completely developed

5.3.2 Test results and user feedback

Descriptive statistics were calculated for the technology acceptance questions, derived from UTAUT (Venkatesh et al., 2012). The results are displayed in Figure 5.27. Ten out of fourteen participants found (somewhat) agreed to the statement that the application would be useful in daily life, which is considered a predictor of eventual usage of the application (Venkatesh et al., 2012).

This is a promising first impression of eventual usage by the pilot users. Furthermore, the test results revealed that the (small) majority of the participants estimated that the application would be easy to use, easy to learn, and easy to become skillful at. The effort expectancy item on clear and understandable interaction with the app received mixed results, which could be explained by the user's limited exposure to the app and not all functionalities being yet implemented. Consequently, users could e.g. not browse through the months in the battery overview and impact overview, which could have caused confusion.

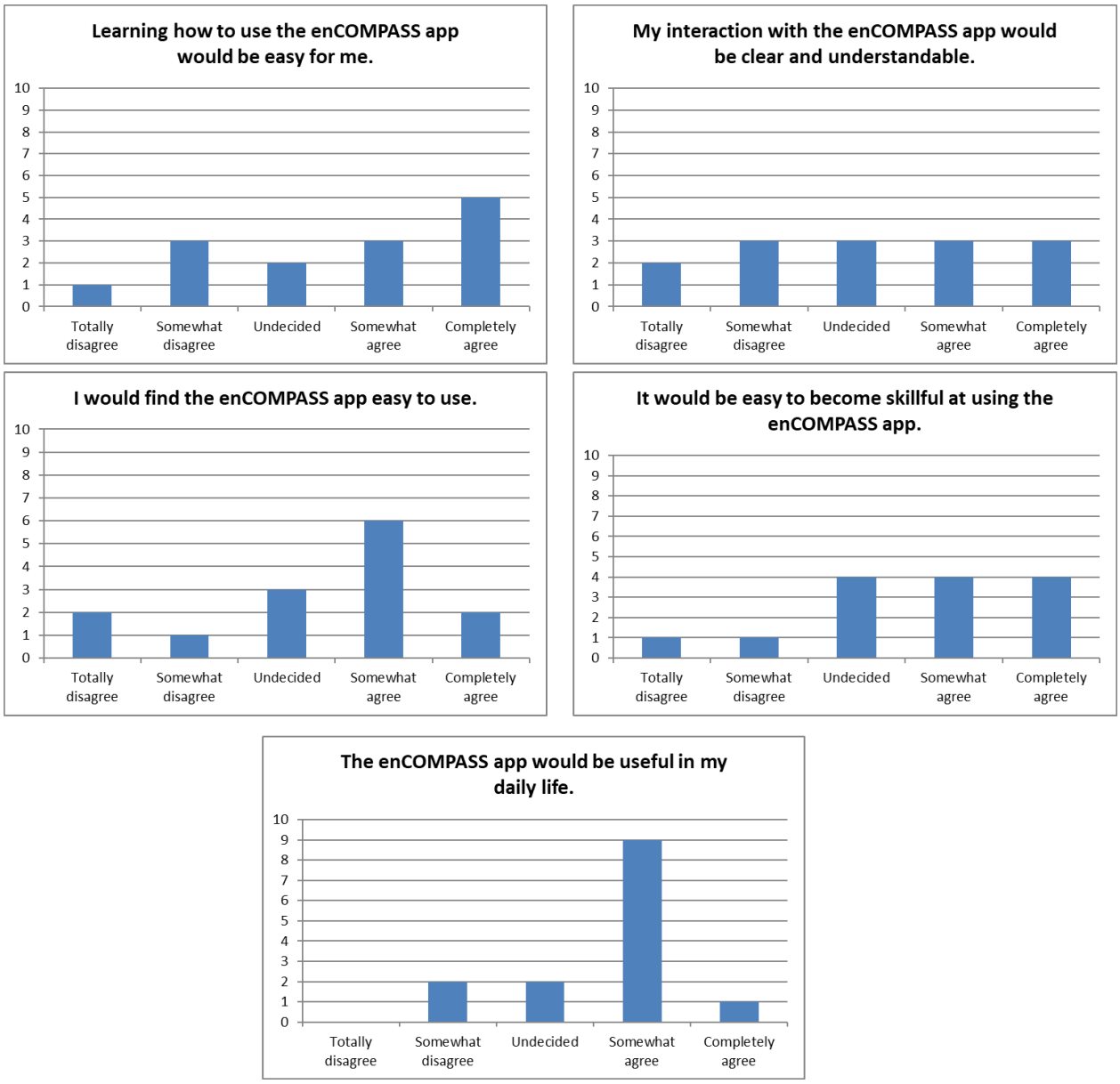


Figure 5.27 Technology acceptance results (Effort expectancy)

Subsequently, the results regarding the motivational impact of the app's features was evaluated. As can be seen from Figure 5.29, participants positively evaluated the motivational impact of the main behavioral change incentives related to consumption feedback. More specifically, the majority of the participants (> 9 out of 14) estimated a positive impact of the development of consumption information over time (e.g. the detailed chart), comparison against a historic baseline (e.g. detailed chart, and battery), comparison against a concrete goal (e.g. battery), and insight into the impact of one's savings on the environment and on the energy bill were found (the impact views).

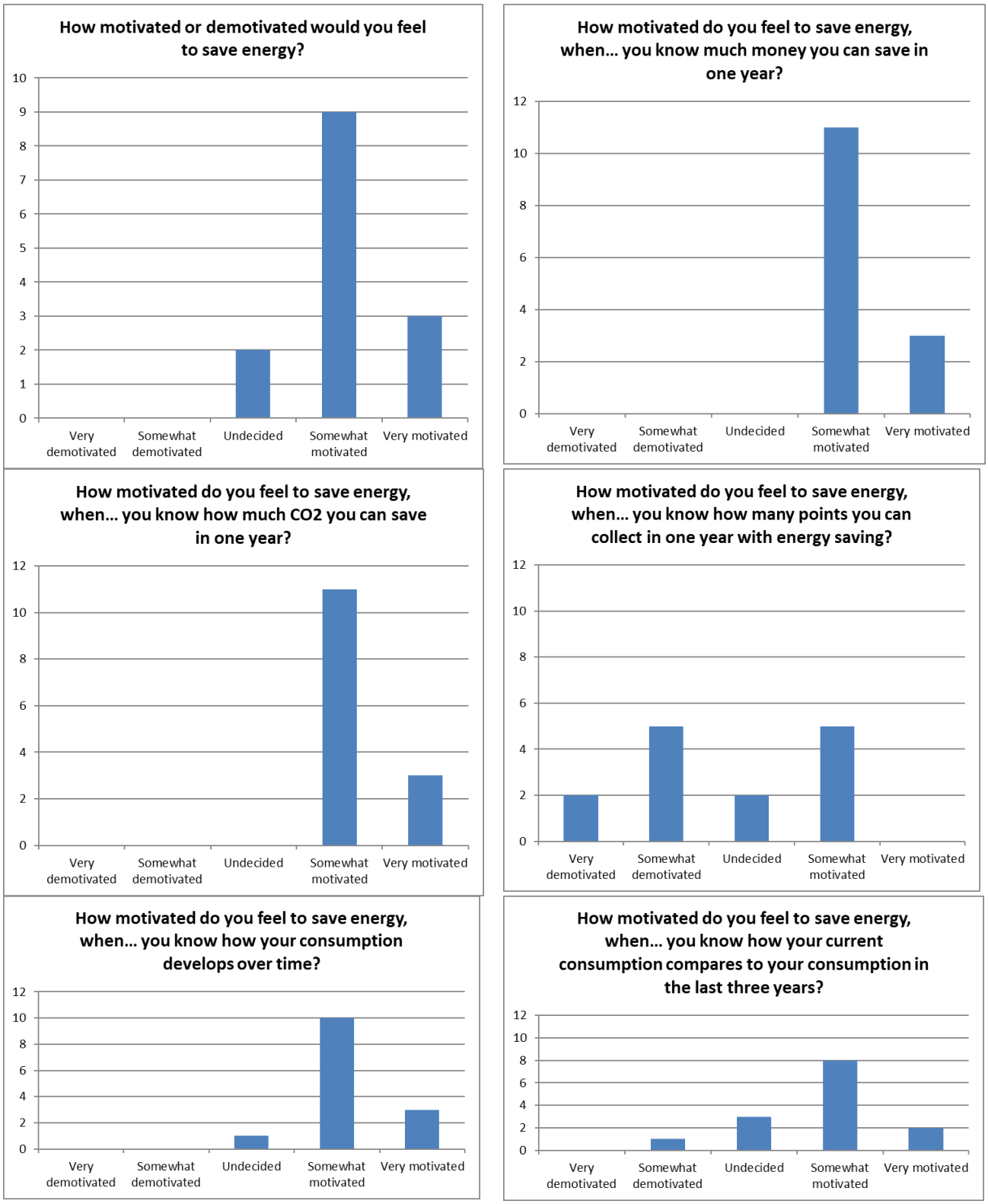


Figure 5.28 Perceived motivational impact of consumption feedback

Consensus among participants was also found concerning the motivational impact of the energy saving tips. Twelve out of fourteen participants felt somewhat or very motivated when they would receive concrete tips to save energy (see Figure 5.30).

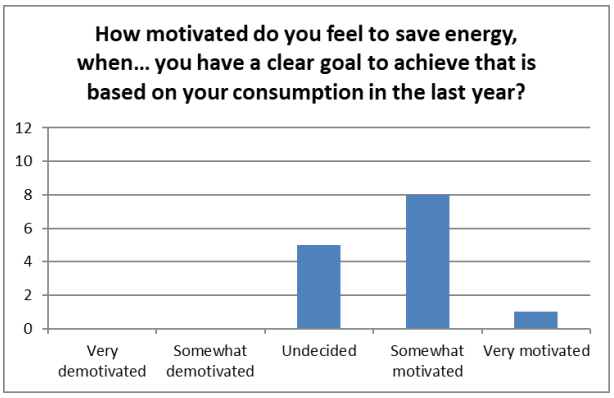


Figure 5.29 Impact of visualizations on energy saving motivation

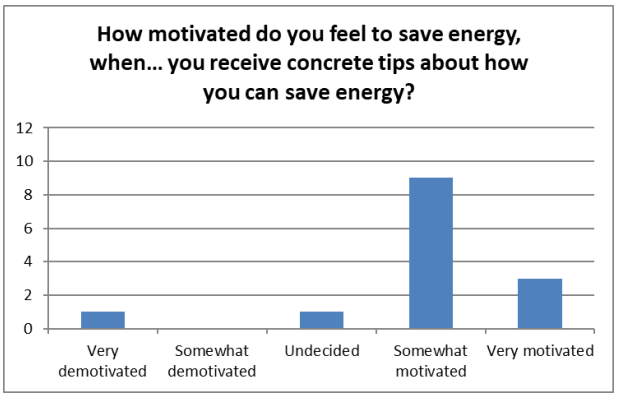
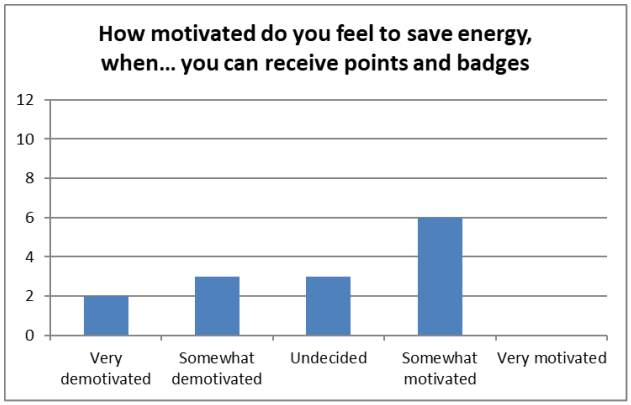
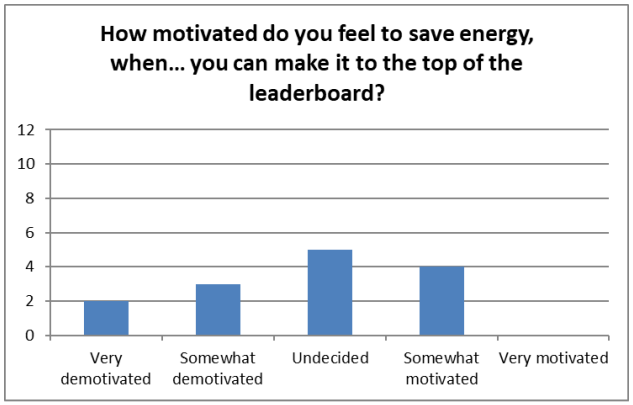


Figure 5.30 Impact of actionable tips on energy saving motivation

As can be seen from Figure 5.31, participants were less positive about the motivational impact of the game elements. However, such results were expected since not all functionalities were fully implemented, and users used the app only for a limited amount of time. As a result, for them it was difficult to get e.g. a sense of competition, or achievement, which would be the main motivational affordances embedded in the game elements. In spite of this this limited exposure and stage of development, still four out of 14 felt motivated by tangible rewards, and by making it to the top of the leaderboard.





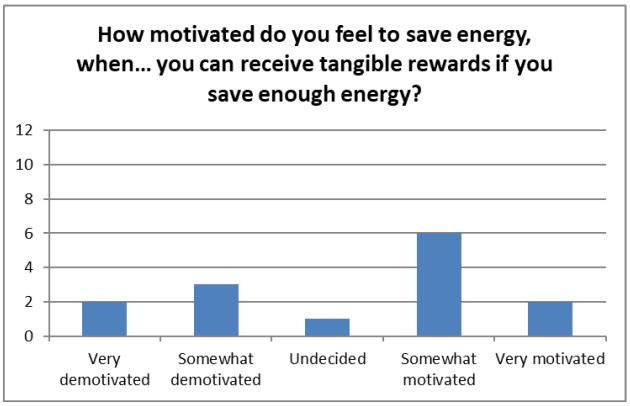


Figure 5.31 Impact of gamification elements on energy saving motivation

Subsequently, the perception of the navigation, and the comprehension of the visualizations were evaluated. Eleven out of the fourteen participants found the application easy to navigate, which suggest

that the main navigation structure is in line with the user's expectations. For the impact visualizations, comprehension of the views regarding monetary and environmental savings was rather good, while in contrast mixed results were found for the battery overview and the having fun view.

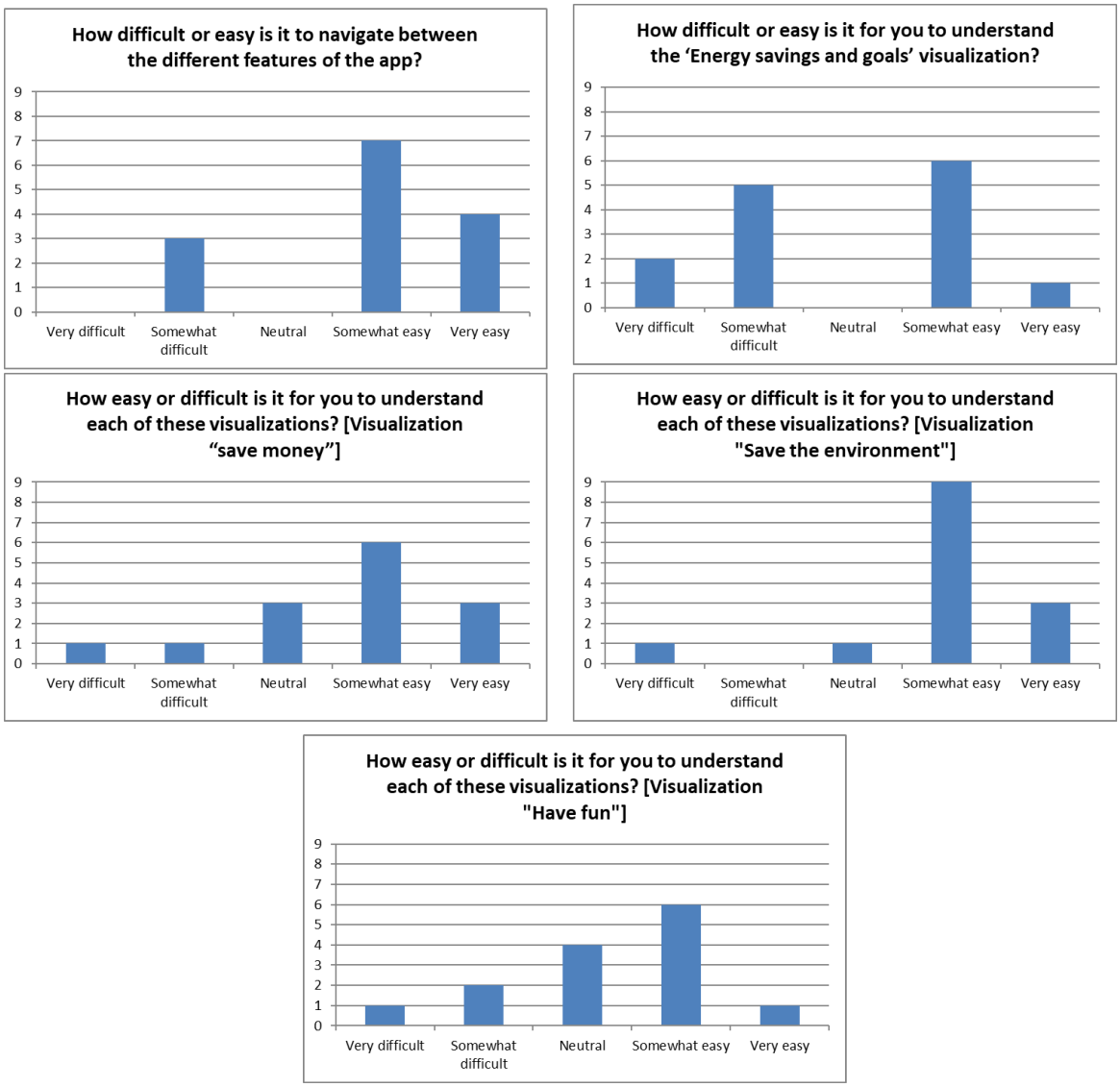


Figure 5.32 Comprehension of the visualizations

Similar to the motivational impact of the game elements, the fun view can only be well understood if the game concept is fully implemented, and users can observe the effect of collecting points, and rewards for a prolonged period of time. This includes an understanding of the type of badges for the thematic areas of saving energy.

Comprehension of the battery overview was also mixed, with half of the participants finding it (somewhat) easy, and the other half finding it (somewhat) difficult to understand. While the impact views contain an at-a-glance view on the total savings and what they mean, the interactive battery overview can only be fully understood when the user can observe what happens when too much energy is consumed, or a goal can or cannot be achieved. In the present test, the battery overview was not fully implemented, with only a single snapshot was embedded. Thus, users could not go back and forth between different months, and different energy achievements, which makes it difficult to fully understand the concept. Qualitative feedback supports this interpretation. E.g. one participant commented: *"The concept of the battery that is running*

down by using more and more energy would be more easy to understand if it's colours were green and red.". In fact, the users in the crowd-test how had more interactive ways of understanding the battery visualization (an animated gif showing a typical interaction sequence) had no such comprehension problems and provided a high assessment of the understandability of the concept. As demonstrated in Section 2, the color-coded normative messages are already part of the concept. However, the snapshot present in the current version does not enable navigating between months and observing the different colors with their meaning. Three other participants mentioned the overview could be improved in the open question on opportunities for improvement.

In the open question concerning the features that participants liked best, the environmental and monetary impact of the user's savings were mentioned (6x), the tips (2x), the battery overview (1x), and the user interface (2x) were mentioned. Some of the comments were ambiguous (e.g. 'the time dimension', 'the section with energy consumptions'). The open question on features that could be improved primarily yielded detailed user interface improvements, save from one participant found many elements 'not so intuitive', and one participant who commented on the understandability of individual pages. E.g. the meaning of the grey color in the battery, a different structuring of the impact views on a menu tab rather than a dropdown, a user having difficulty navigating between the tips, and a user commenting on the 'delicacy' of manipulating the slider to adjust the time span in the detailed consumption chart.

Two participants commented on part of the menu bar not being completely visible on their Samsung Galaxy phones. Participants offered several concrete suggestions for these user interface improvement, which will be considered in the upcoming sprints towards the first release.

One participant did not understand the rewards section, which can be understood since currently few actions yield points, the page does not yet contain an explanation for when users are eligible for rewards, and the rewards themselves are not yet finalized. Such details will be clarified before the release of the app to the pilot users.

In conclusion, in spite of the very early stage of development at which the test was conducted, participants were positive about the concept behind the consumption feedback views, the goal setting, and the tips, while the game elements were less appreciated, presumably due to limited exposure, and the features not being fully implemented.

6 CONCLUSION

This deliverable has iteratively designed and evaluated the visualization concept employed in enCOMPASS, which comprises different views on energy consumption, whereby in addition to the commonly used visualizations in e.g. bar charts that for a class of users are difficult to understand, metaphors were employed, for the purpose of a) monitoring consumption and comparing one's consumption level against a clear goal, and against one's historic consumption, and b) raising awareness about the impact of one's consumption. This was done by metaphorically demonstrating the impact on the energy bill (e.g. 'Saving money'), the environment ('Protecting the environment'), and the game achievements ('Having fun'), while the default impact view is adapted to the user's energy saving motivation.

Test results of both the crowdsourcing-based tests, and the first test with the integrated awareness app has yielded support for the visualization concepts, in terms of comprehension, and in terms of the perceived impact on the motivation for energy saving.

Useful suggestions for further improvements were received for all visualizations, the majority of which were partly already implemented and displayed in Section 3, and partly require further thought and alignment with technical partners, preceding their implementation before the first release of the platform in M18.

7 REFERENCES

Atkinson, J. W. and Litwin, G. H. (1960). Achievement motive and test anxiety conceived as motive to approach success and motive to avoid failure. *The Journal of Abnormal and Social Psychology, 60*(1), p. 52-63.

Bamberg, S. (2013). Changing environmentally harmful behaviors: A stage model of self-regulated behavioral change. *Journal of Environmental Psychology*, *34*, 151–159.

Fréjus, M., & Martini, D. (2016). Why Energy Consumption Feedback Is not (Only) a Display Issue. In A. Marcus (Ed.), *Lecture Notes in Computer Science. Design, User Experience, and Usability: Novel User Experiences* (Vol. 9747, pp. 461–471).

Froehlich, J., Patel, S., Landay, J. A., Findlater, L., Ostergren, M., Ramanathan, S., … Bai, M. (2012). The design and evaluation of prototype eco-feedback displays for fixture-level water usage data. In *Proceedings of the 2012 ACM annual conference on Human Factors in Computing Systems - CHI '12* (p. 2367). New York, New York, USA: ACM Press.

Gölz, S., & Hahnel, U. J. J. (2016). What motivates people to use energy feedback systems? A multiple goal approach to predict long-term usage behaviour in daily life. *Energy Research and Social Science*, *21*, 155–166. https://doi.org/10.1016/j.erss.2016.07.006

Gustafsson, A., Bång, M., & Svahn, M. (2009). Power explorer: a casual game style for encouraging long term behavior change among teenagers. *ACE '09 Proceedings of the International Conference on Advances in Computer Entertainment Technology*, 182–189.

Hassenzahl, M. (2004). The interplay of beauty, goodness, and usability in interactive products. *Human-computer interaction*, *19*(4), 319-349.

Jacucci, G., Spagnolli, A., Gamberini, L., Chalambalakis, A., Björksog, C., Bertoncini, M., … Monti, P. (2009). Designing effective feedback of electricity consumption for mobile user interfaces. *PsychNology Journal*, *7*(3), 265–289.

Karjalainen, S. (2011). Consumer preferences for feedback on household electricity consumption. *Energy and Buildings*, *43*(2–3), 458–467.

Karlin, B., Zinger, J. F., & Ford, R. (2015). The Effects of Feedback on Energy Conservation: A Meta-Analysis. *Psychological Bulletin*, *141*(6), 1205–1227.

Lindenberg, S., & Steg, L. (2007). Normative, Gain and Hedonic Goal Frames Guiding Environmental Behavior. *Journal of Social Issues*, *63*(1), 117–137.

Micheel, I., Novak, J., Fraternali, P., Baroffio, G., Castelletti, A., & Rizzoli, A.-E. (2015). Visualizing & Gamifying Water & Energy Consumption for Behavior Change. In *IFIP INTERACT 2015 Adjunct Proceedings*.

Monigatti, P., Apperley, M., & Rogers, B. (2010). Power and energy visualization for the micro-management of household electricity consumption. *Proceedings of the International Conference on Advanced Visual Interfaces - AVI '10*, 325.

Novak, J., Melenhorst, M., Micheel, I., Pasini, C., Fraternali, P., Rizzoli, A.E. (2018). Integrating behavioural change and gamified incentive modelling for stimulating water saving. *Environmental Modelling and Software* (article in press).

Oinas-Kukkonen, H. (2013). A foundation for the study of behavior change support systems. *Personal and Ubiquitous Computing*, *17*(6), 1223–1235.

Rist, T. (2014). Towards a More Responsible Use of Energy through Visualization of Energy Data, 9–14.

Schwartz, T., Stevens, G., Jakobi, T., Denef, S., Ramirez, L., Wulf, V., & Randall, D. (2015). What people do with consumption feedback: A long-term living lab study of a home energy management system. *Interacting with Computers*, *27*(6), 551–576.

Spagnolli, A., Corradi, N., Gamberini, L., Hoggan, E., Jacucci, G., Katzeff, C., … Jonsson, L. (2011). Eco-feedback on the go: Motivating energy awareness. *Computer*, *44*(5), 38–45.

Sundramoorthy, V., Liu, Q., Cooper, G., Linge, N., & Cooper, J. (2010). DEHEMS: A user-driven domestic energy monitoring system. *2010 Internet of Things, IoT 2010*, (January).

Steg, L., Perlaviciute, G., Van der Werff, E., & Lurvink, J. (2013). The significance of hedonic values for environmentally-relevant attitudes, preferences and actions. Environment and Behavior, DOI: 10.1177/0013916512454730

Schwartz, T., Stevens, G., Jakobi, T., Denef, S., Ramirez, L., Wulf, V., & Randall, D. (2015). What people do with consumption feedback: A long-term living lab study of a home energy management system. *Interacting with Computers*, *27*(6), 551–576.

Steg, L., Bolderdijk, J. W., Keizer, K., & Perlaviciute, G. (2014). An Integrated Framework for Encouraging Pro-environmental Behaviour: The role of values, situational factors and goals. *Journal of Environmental Psychology*, *38*, 104–115.

Tiefenbeck, V. (2016). Long-lasting effects or short-term spark? On the persistence of behaviour change induced by real-time feedback on resource consumption. In *24th European Conference on Information Systems (ECIS 2016)* (Vol. 84).

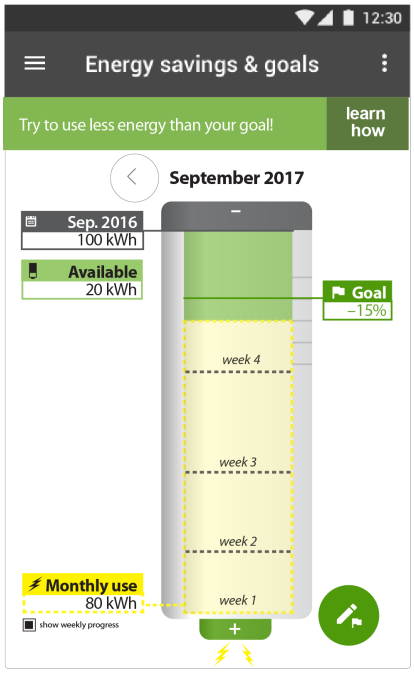
Venkatesh, V., Thong, J. Y. L., Xin, X., & Xu, X. (2012). Consumer Acceptance and Use of Information Technology: Extending the Unified Theory of Acceptance and Use of Technology. *MIS Quarterly*, *36*(1), 157–178.

Voss, K. E., Spangenberg, E. R., & Grohmann, B. (2003). Measuring the hedonic and utilitarian dimensions of consumer attitude. *Journal of marketing research*, *40*(3), 310-320.

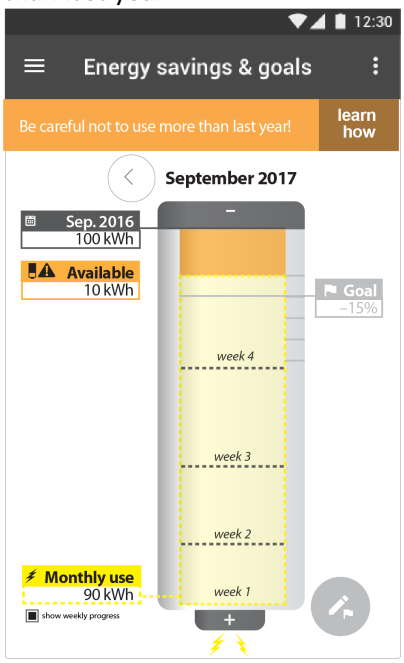
8 APPENDICES

Appendix 1 Visualizations used in the DIRECTION 2 treatment.

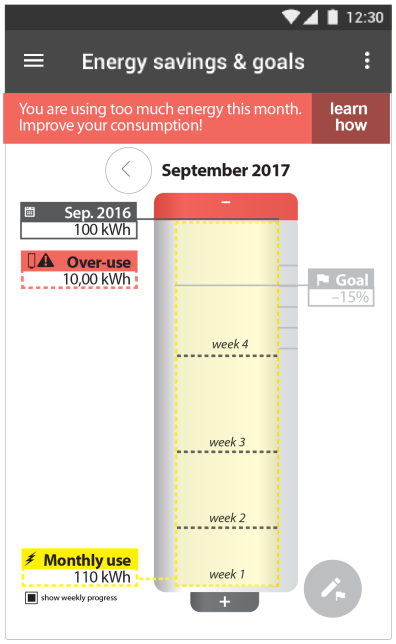
Green: the user is on track to meet their goal



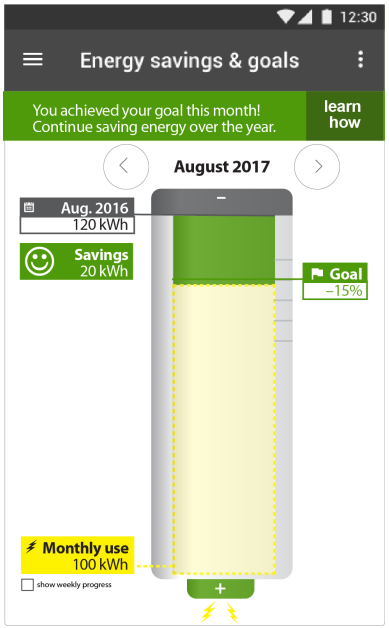
Orange: the user has surpassed their goal, but has not used more than last year



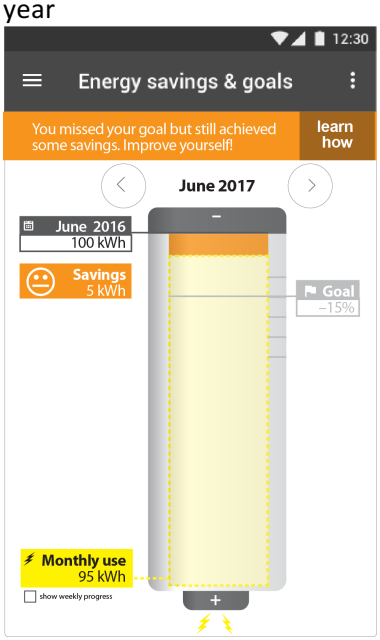
Red: the user has used more than last year



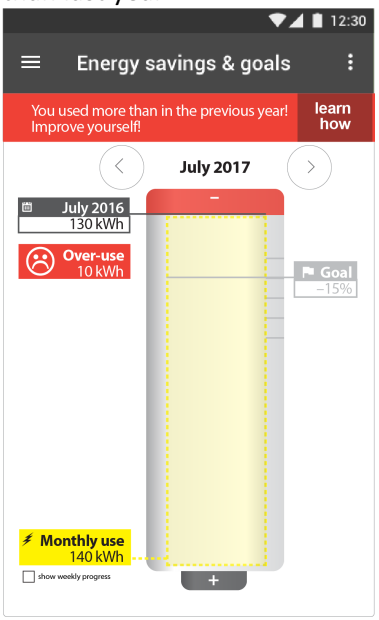
Green: the user has achieved the goal



Orange: the user has not reached their goal, but used less than last year



Red: the user has not achieved the goal and also used more than last year



Appendix 2 Full Questionnaire used in the DIRECTION 1 and DIRECTION 2 Overview visualization test[7].

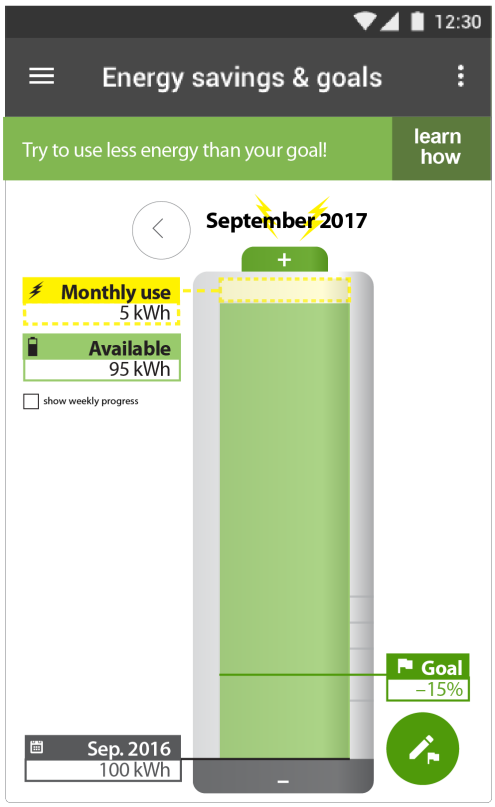
The Scenario: The Energy Saving Application

Imagine you are using an application on your smartphone that helps you to save energy. It gives you tips for energy saving, shows how much energy you are using, and also how much energy you have already saved. The feature we would like to test is monitoring your monthly savings. The idea is to spend less energy compared to last year. You will set a target savings goal for the current month. The application will then allow you to monitor this month's consumption and provide you with feedback on your savings.

In the following we will show you some screenshots of the application for the months up to September. We will also ask ask you a series of questions. Let's imagine that this September you decide to **set your savings goal to use 15% less energy** compared to September last year.

PAGE 1: First Week of September

Imagine that you have just started using the application. *Please answer the questions below.*



[7] - Here we present an example with the DIRECTION 1 visualization. The survey used for the DIRECTION 2 visualization was identical. The images used can be seen in Section 2.3

Question 1. Using the information from the screenshot above, for each of the statements below, please indicate whether they are True or False

	TRUE	FALSE
1.1 I have already reached my savings goal	○	○
1.2 This month I can not use more than 100kWh of electricity	○	○
1.3 This month, I would like to use 85kWh of electricity	○	○

Question 2. What do you think will happen if you continue to use electricity this month? How would this be reflected in the battery?

Please write at least two full sentences. The sentences should be logically and grammatically correct. Please refer in your answer to specific image elements.

Question 3. What do you think will happen if you click on the green button in the bottom right of the screen?

Please write at least one full sentence. The sentence should be logically and grammatically correct. Please refer in your answer to specific image elements.

Question 4. In your opinion, is this visualization clear?

Please choose one answer option that best matches your opinion.

	1	2	3	4	
Confusing	○	○	○	○	○ Clear

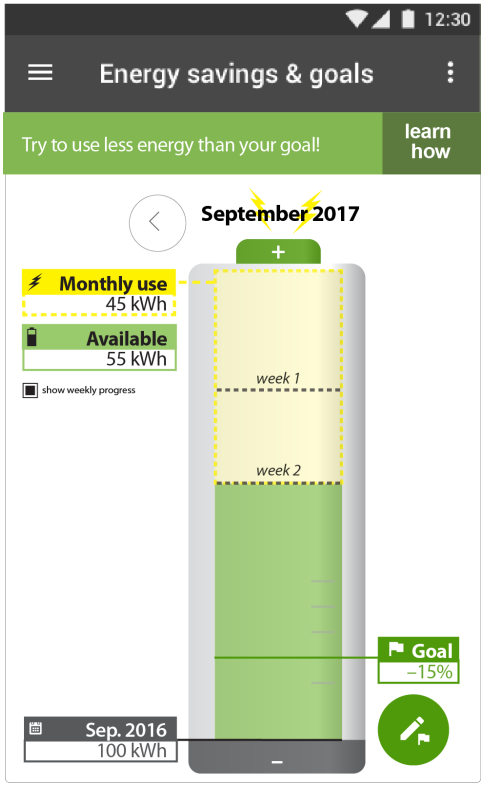
Question 5. Please explain your answer to the previous question. Which elements in this screenshot are confusing? Which elements are clear? Why?

Please answer in at least two complete sentences. Please refer in your answer to specific image elements.

PAGE 2: Third Week of September

Imagine that two weeks have passed and you obtain the following screenshot in the application.

Please answer the questions below.



Question 6. Using the information from the screenshot, for each of the statements below, please indicate whether they are True or False

	TRUE	FALSE
6.1 Over the two weeks, I have used 45kWh of electricity	○	○
6.2 In week 1 I have used more electricity than in week 2	○	○

7. How likely are you to meet your goal if you continue to use the same amount of electricity over the next two weeks? Please explain.

Please write at least two full sentences. The sentences should be logically and grammatically correct.

---

8. What do you think the plus and minus signs in the battery mean?

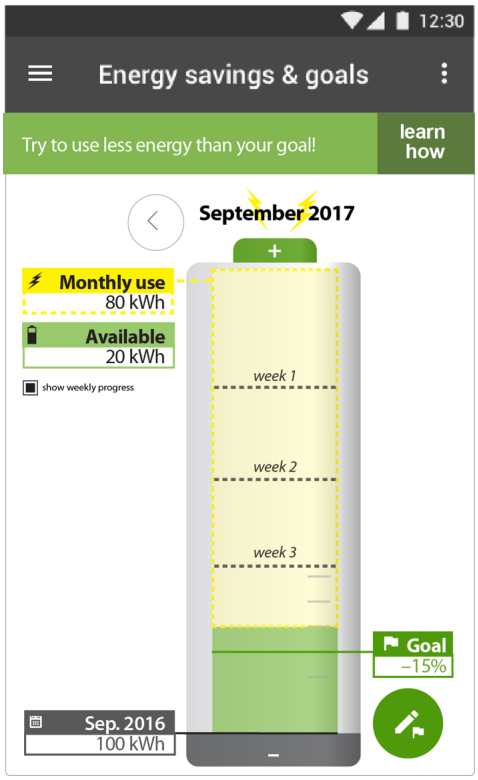
Please explain in one complete sentence. The sentence should be logically and grammatically correct.

PAGE 3: The fourth week of September

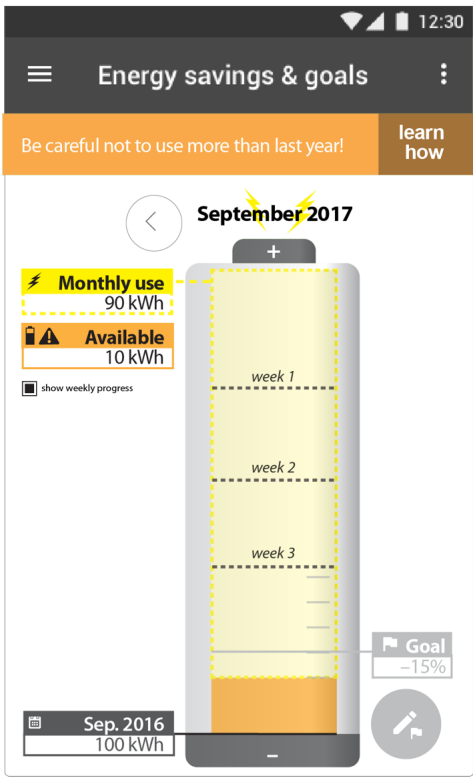
Another week has passed - you are now in the fourth week of September. Two scenarios could happen at this point, depending on your consumption. **Scenario 1 and Scenario 2. Both scenarios are equally likely.**

Please answer the questions below.

Scenario 1



Scenario 2



9. In which of the two scenarios are you on the right track to meet your savings goal for September?

Please choose one answer option.

- ○ Scenario 1
- ○ Scenario 2
- ○ Both scenario 1 and 2
- ○ Neither scenario 1 nor 2

10. What do you think the orange color means in Scenario 2?

Please explain in one complete sentence. The sentence should be logically and grammatically correct.

---

11. What do you think the green color means in Scenario 1?
Please explain in one complete sentence. The sentence should be logically and grammatically correct.

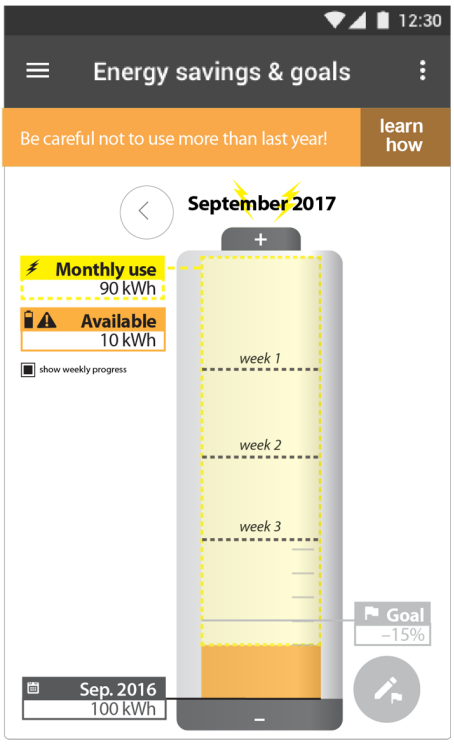
---

12. In scenario 2, why do you think the goal setting button turns grey?
Please answer in one complete sentence. The sentence should be logically and grammatically correct.

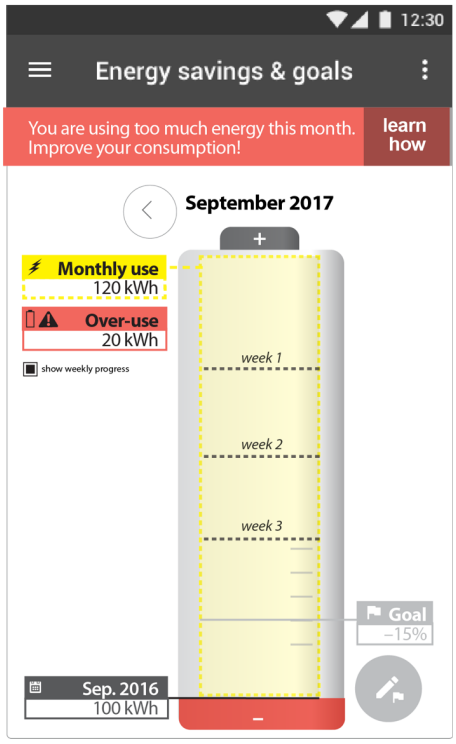
PAGE 4.

Now imagine that certain events have led to Scenario 3. You are asked to compare Scenario 2 that you have seen on the previous page and Scenario 3.
Please answer the questions below.

Scenario 2



Scenario 3



13. In your opinion, what is the difference between the Scenarios 2 and 3?
Please explain in two complete sentences. The sentences should be logically and grammatically correct.

---

14. What do you think the red color means in Scenario 3?

Please explain in one complete sentence. The sentence should be logically and grammatically correct.

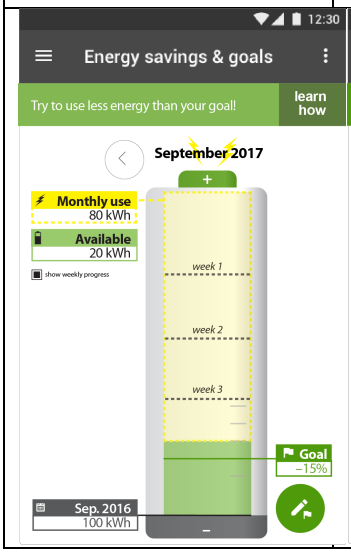
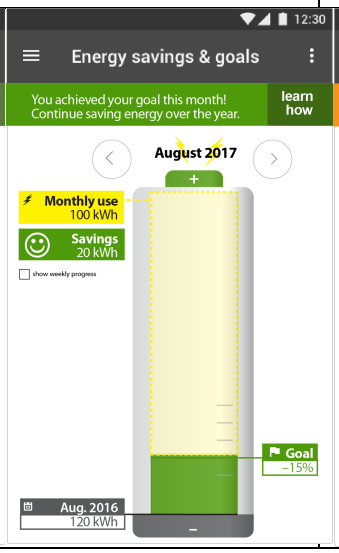
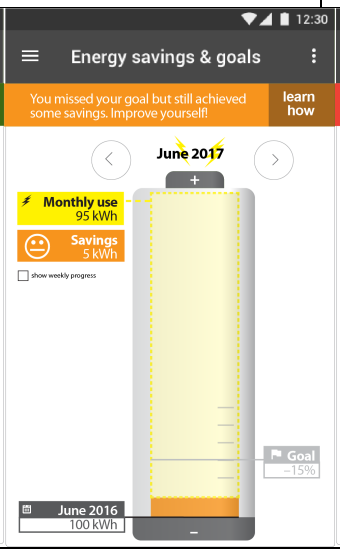
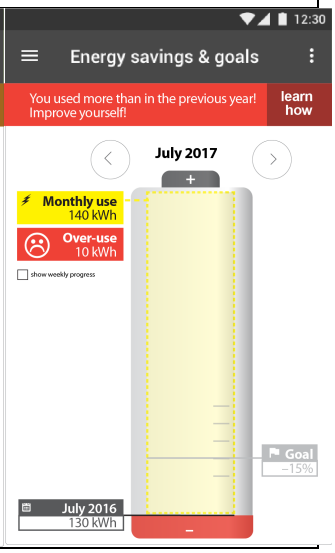
For each of the statements below relating to Scenario 3, please indicate whether they are True or False

	TRUE	FALSE
14.1 In Scenario 3 I can not spend any more energy	○	○
14.2 In Scenario 3 I have spent more than in September 2016	○	○

PAGE 5.

Imagine you are still in the fourth week of September and you are curious whether you can achieve your goal for this month. You decide to check how you have performed in the previous months. To illustrate this, we have prepared an animation.

Please look at the animation and answer the questions below. If you cannot view this animation, it can also be viewed on YouTube. Please open the link in the new tab[8].

Screen 1	Screen 2	Screen 3	Screen 4
			

[8] - As discussed in section 4.1.1. in order to present the dynamics in the application, we have used an animated gif image of the four pictures presented below.

15. In which of the months have you achieved your SAVINGS GOAL?
Please check all that apply. Please note that you can only achieve your savings goal at the end of the month.

- [ ] September
- [ ] August
- [ ] July
- [ ] June
- [ ] I don't know

16. What do you think the smileys mean?
Please explain in one complete sentence.

---

17. What is the difference between the light green September and the bright green August?
Please explain your answer in two complete sentences. The sentences should be logically and grammatically correct.

---

18. In which of the months have you USED LESS ELECTRICITY than in the previous month last year?
Please check all that apply.

- [ ] August
- [ ] July
- [ ] June
- [ ] I don't know

Please explain your answer to the previous question in one complete sentence. The sentence should be logically and grammatically correct.

PAGE 6.

Below you see the same animation from the previous question. It is there just as a reminder. We now ask you to evaluate this visualization of energy savings as a whole.
Please answer the questions below. If you cannot view this animation, it can also be viewed on YouTube. Please open the link in the new tab.(HERE THE SAME VISUALIZATION FROM THE PREVIOUS PAGE)

19. When I see this visualization, I am willing to put extra effort into saving energy.
Please choose one answer option that best represents your opinion.

- ○ Strongly disagree
- ○ Somewhat disagree
- ○ Neither agree nor disagree
- ○ Somewhat agree
- ○ Strongly agree

20. When I see this visualization, I am motivated to meet my monthly savings goal.
Please choose one answer option that best represents your opinion.

- ○ Strongly disagree
- ○ Somewhat disagree
- ○ Neither agree nor disagree
- ○ Somewhat agree
- ○ Strongly agree

21. How do you evaluate this visualization overall?
Please choose one answer option for each of the categories.

	1	2	3	4	5	
lame	○	○	○	○	○	exciting
challenging	○	○	○	○	○	easy
complicated	○	○	○	○	○	simple
confusing	○	○	○	○	○	clear

not fun	○	○	○	○	○	fun
ugly	○	○	○	○	○	beautiful
ineffective	○	○	○	○	○	effective

22. What is it that you LIKE about this visualization?
Please mention at least two things. The sentences should be logically and grammatically correct and relate to the image above.

---

23. What is it that you DO NOT LIKE about this visualization?
Please mention at least two things. The sentences should be logically and grammatically correct, relate to the image above and differ from what you wrote in the previous question.

---

24. On a scale from 0 - 10, how do you evaluate this visualization overall? (0 - very low; 10 - very high)
Please select one answer that matches your opinion the most.

Answer options from 0 – 10.

PAGE 7. Demographics.

25. Where are you from?

(LIST of all countries in EU, plus UK, Russia, Ukraine)

26. How old are you?

- ○ below 18 years old
- ○ 18-24 years old
- ○ 25-34 years old
- ○ 35-44 years old
- ○ 45-54 years old
- ○ 55-64 years old

- ○ 65-74 years old
- ○ 75 years or older
- ○ I'd rather not say

27. What is your gender?

- ○ Male
- ○ Female
- ○ Other / I'd rather not say

28. How many people are in your household?

- ○ 1
- ○ 2
- ○ 3
- ○ 4
- ○ 5 or more

29. What is your highest level of completed education?

- ○ Middle School
- ○ High School
- ○ Vocational education
- ○ Bachelor Degree
- ○ Master Degree
- ○ PhD
- ○ None of these

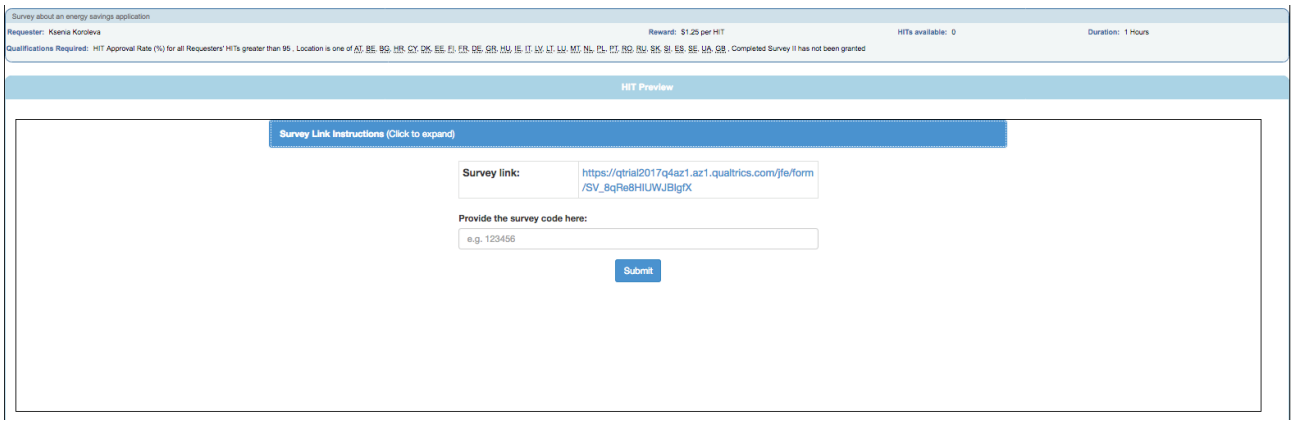
Your feedback *(Optional)*

We would appreciate any remarks, suggestions, or other thoughts about the HIT, or the visualizations. This would help our research and the design of the next HITs we plan to run.

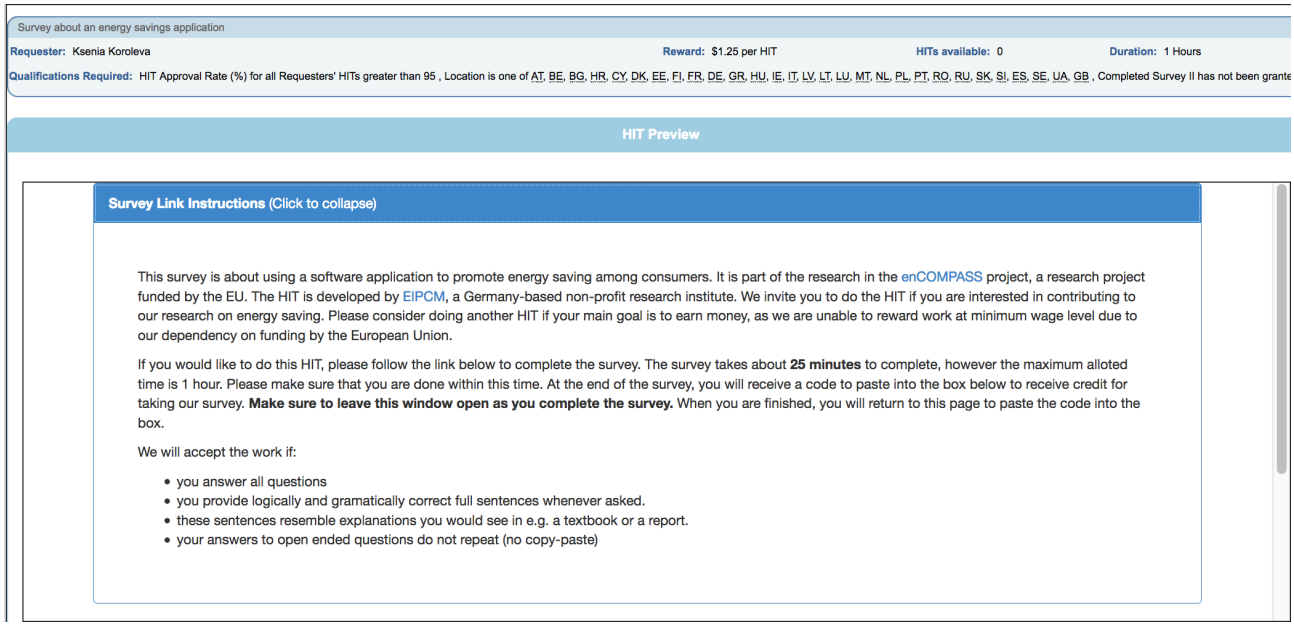
30. Your feedback, thoughts and suggestions.

Appendix 3 Screenshot of the Mechanical Turk Assignment (for the Overview Visualization, DIRECTION 1 version)

- The shortened view -



- The Expanded View of the instructions -



Appendix 4 The distribution of the answers to understandability concept of the participants in the Direction 2 Treatment.

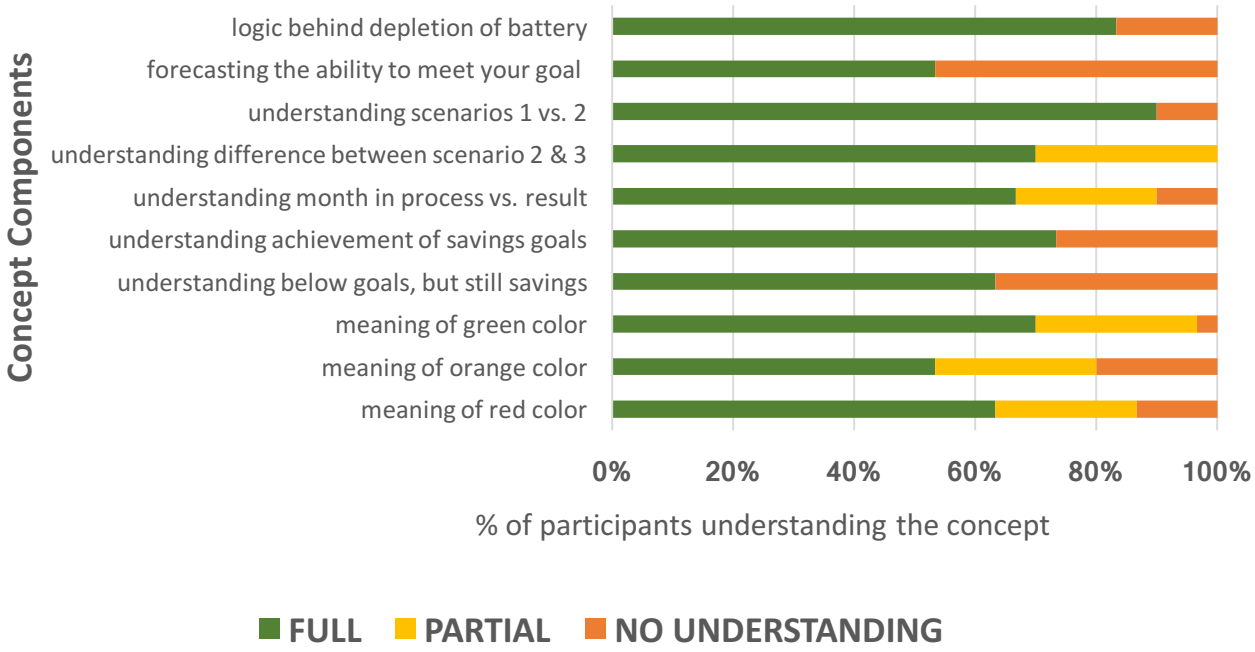


Figure 8.1 Concept Understandability by the participants in the Direction 2 Treatment.

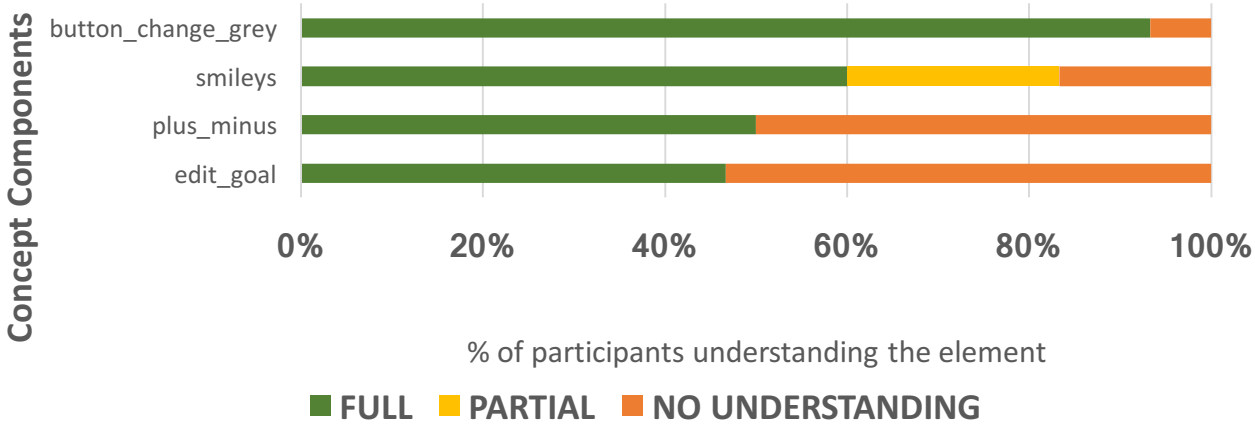


Figure 8.2 Understandability of the UI elements by the participants in the Direction 2 Treatment

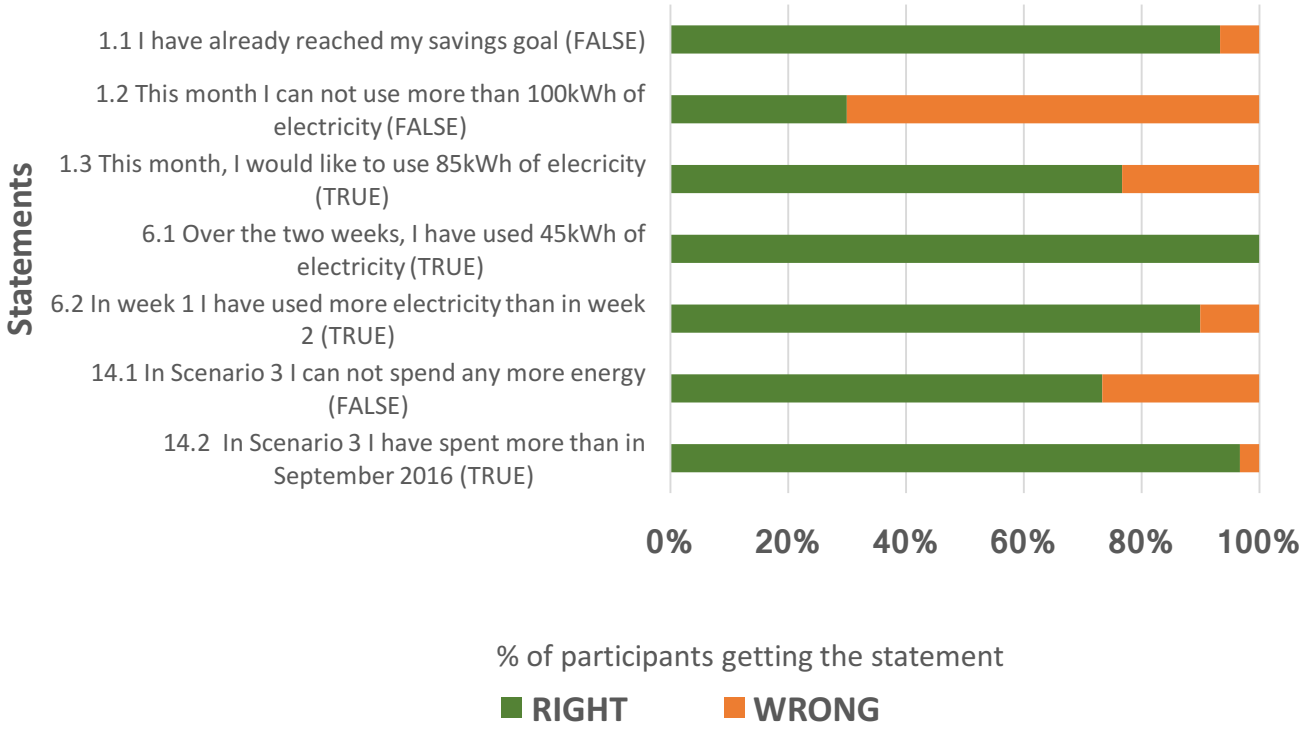


Figure 8.3. Distribution of answers to the True/False statements in the Direction 2 treatment.

Appendix 5 Means of all evaluation items and mean differences between treatments

	treatment	N	Mean	Std. Deviation	Std. Error Mean	Mean difference test
Evaluation of confusing-clear in the beginning of the survey	direction 1	31	3.58	1.148	.206	t(59) = 1.391, p = .169
	direction 2	30	3.13	1.358	.248	
Evaluation of lame-exciting (hedonic quality 1)	direction 1	31	3.45	.961	.173	t(59) = 0.481, p = .632
	direction 2	30	3.33	.959	.175	
Evaluation of challenging-easy (hedonic quality 2)	direction 1	31	3.74	.930	.167	t(55) = 0.154, p = .878
	direction 2	30	3.70	1.179	.215	
Evaluation of complicated-simple (Pragmatic quality 1)	direction 1	31	3.81	1.108	.199	t(59) = .684, p = .497
	direction 2	30	3.60	1.248	.228	
Evaluation of confusing-clear in the end of the survey (pragmatic quality 2)	direction 1	31	3.81	1.138	.204	t(59) = 1.038, p = .303
	direction 2	30	3.50	1.167	.213	
Hedonic Attitude (fun – not fun)	direction 1	31	3.32	1.013	.182	t(54.6) = -.480, p = .631
	direction 2	30	3.47	1.306	.238	
Beauty perception (ugly-beautiful)	direction 1	31	3.39	1.022	.184	t(59) = -.051, p = .960
	direction 2	30	3.40	.968	.177	
Utilitarian attitude (ineffective-effective)	direction 1	31	4.00	.894	.161	t(59) = .289, p = .774
	direction 2	30	3.93	.907	.166	

Appendix 6 Questionnaire used in the Impact visualization test[9].

Survey about the Energy Saving Application

This survey is about using a software application to promote energy saving among consumers. It is part of the research in the enCOMPASS project, a research project funded by the EU conducted by EIPCM, a Germany-based non-profit research institute.

This survey consists of several parts. You will be kindly asked to provide information about what you find important in life and to evaluate two versions of the energy saving application that we have developed.

The survey will take about 25 minutes to complete. Please complete the survey without interruptions. Thank you!

[9] - The Impact questionnaire was conducted two times: one for the Monetary/Environmental Visualization (test 1) and one for the Hedonic visualization (test 2). Here we present the questions asked for the monetary visualization. The questionnaire asked exactly the same questions, but with a different set of pictures for the other two visualizations: environmental and hedonic. Environmental questions were asked right after the Monetary visualization to the same participants. The questions for the hedonic visualization were asked separately (during test 2). We also present these questions in this test with a special asterisk*.

Question 1. Below you will find 12 things people find important in life. Please rate how important each of these things is for you AS A GUIDING PRINCIPLE IN YOUR LIFE.

Please choose one answer option for each category.

	Opposed to my principles	Not important at all	Low importance	Slightly important	Moderately important	Very important	Extremely important
Harmony with other species	○	○	○	○	○	○	○
Fitting into nature	○	○	○	○	○	○	○
Preserving nature	○	○	○	○	○	○	○
Preventing pollution	○	○	○	○	○	○	○
Control over others, dominance	○	○	○	○	○	○	○
Material possessions, money	○	○	○	○	○	○	○
The right to lead or command	○	○	○	○	○	○	○
Having an impact on other people and events	○	○	○	○	○	○	○
Hard working, aspiring	○	○	○	○	○	○	○
Joy, gratification of desires	○	○	○	○	○	○	○
Enjoying life: enjoying food, sex, leisure, etc.	○	○	○	○	○	○	○
Doing pleasant things	○	○	○	○	○	○	○

Question 1a^[10]. People save energy for different reasons. Two of them are saving on the energy bill (monetary saving) and reducing the impact on the environment. Which benefit is more important to you personally?

Please choose the answer that represents your true opinion.

	1	2	3	4	5	6		
Environmental Impact	○	○	○	○	○	○	○	Monetary Savings

The Scenario: the Energy Saving Application

Imagine you are using an application on your smartphone that helps you to save energy. It gives you tips for energy saving, shows how much energy you are using, and also how much energy you have already saved. Right now, we are testing a new feature which helps you understand the impact of your energy saving efforts. We have developed two versions of this feature. Each version displays the impact of your energy savings in a different way. In the following, we will ask you some questions about each of these versions.

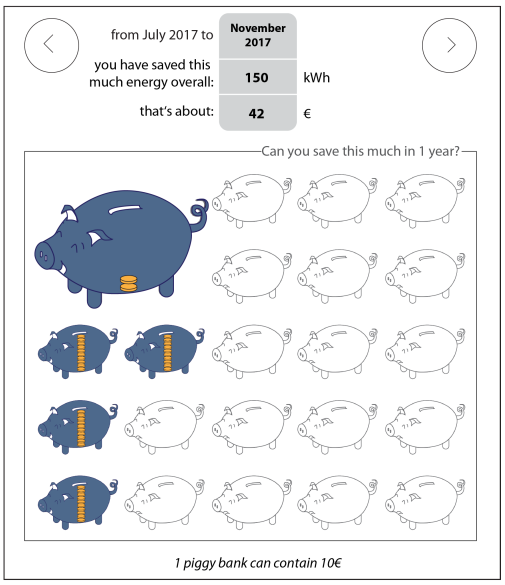
The Scenario: the Energy Saving Application *

Imagine you are using an application on your smartphone that helps you to save energy. It gives you tips for energy saving, shows how much energy you are using, and how much energy you have already saved. The application is designed as a game: you can collect points, earn badges, and win the leaderboard competition. You can get points and badges for saving energy and for regularly using the app (e.g. read tips, check out your consumption, and so on). Right now, we are testing a new feature which helps you understand the impact of your energy saving efforts. It's designed as a challenge: how much energy can you save? Each month, you set an energy saving goal for yourself. If you meet the energy saving goal, you are awarded points. In the following we would like to test the visualization that we have developed for displaying the points you have earned.

[10] - This and all other questions marked with ^ were only asked during test 1.

PAGE 1. This is the energy saving application.

This is an **example** screenshot.
Please answer the questions below.



Question 2. Please name two main graphical elements in this screenshot.
Please fill in one word in each blank. The word should be a known word in English and relate to the graphical elements in the screenshot.

- Graphical Element 1 _____
- Graphical Element 2 _____

Question 3. In your opinion, what do the graphical elements symbolize in relation to energy savings in this screenshot?
Please explain in 2 complete sentences. The sentences should be logically and grammatically correct.

_____

_____

Question 3a.* In your opinion, what do the different colors mean in this screenshot?
Please explain in one complete sentence. The sentence should be logically and grammatically correct.

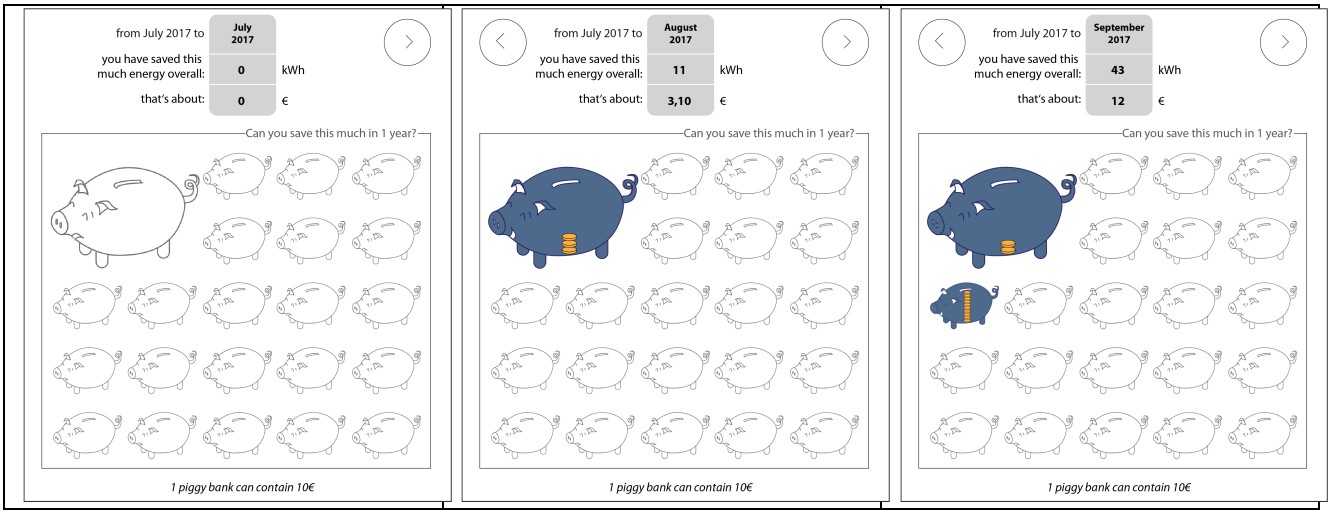
_____

_____

PAGE 2. Now imagine you start using the energy saving application.

This animation illustrates your usage of the energy saving application between July and September.

After seeing this animation please answer the questions below.[11]



Question 4. What is the logic behind appearing and disappearing piggy banks in the picture?

Please explain in two complete sentences. The sentences should be logically and grammatically correct.

_____

_____

Question 5. Please determine whether each of these statements is TRUE or FALSE based on the information that you can obtain from the animation.

	TRUE	FALSE
1. I have saved nothing in July	○	○
2. Out of the three months, I have saved the most in September	○	○
3. I have saved more in August than in September	○	○

[11] - This set of illustrations was presented to the user in the form of animated gif, so that they could see some dynamics.

Question 6. When answering the questions above, I used the following information:
Please check all that apply

- ☐ Graphics
- ☐ Textual information
- ☐ Amont of KwH
- ☐ Amount of EUR
- ☐ None of the above

Question 4a.* In your opinion, what is the connection between energy savings and the beginner, intermediate and expert saver symbols?
Please explain in one complete sentence. This sentence should be logically and grammatically correct.

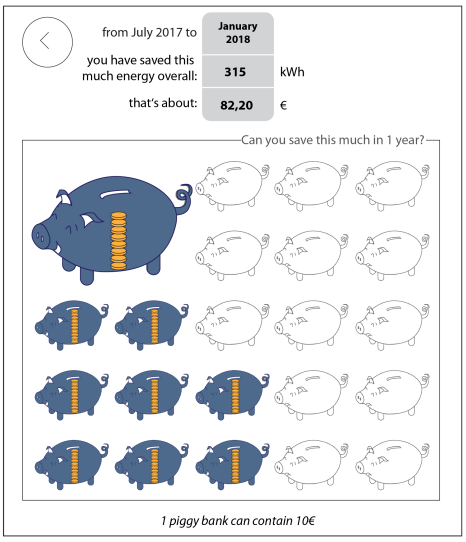
_____

_____

Question 4b.* What kind of saver are you by the end of September?
Please choose the answer option that applies.

- ○ Beginner Saver
- ○ Intermediate Saver
- ○ Expert Saver
- ○ None of the above

PAGE 3. Imagine you continue using the application and **after half a YEAR** you obtain the following screenshot.
Please answer the question below.



Question 7. How much have you saved overall so far compared to what you could potentially save?
Please choose one answer option. You might be asked to explain your answer in the next question.

- ○ I have saved nothing yet.
- ○ I have saved just a little, but could save much more.
- ○ I have saved quite a lot, but could save some more.
- ○ I have saved all I could already.
- ○ I don't know how much I have saved.

Question 7a.* Which kind of saver are you by the end of January 2018?
Please choose the answer option that applies.

- ○ Beginner Saver
- ○ Intermediate Saver
- ○ Expert Saver
- ○ None of the above

PAGE 4. Now please evaluate the energy saving application[12].

Question 8. When I see this visualization, I am willing to put extra effort into saving energy.
Please choose one answer option that best represents your opinion.

[12] - here the same picture as on the previous page was shown to the user.

- ○ Strongly disagree
- ○ Somewhat disagree
- ○ Neither agree nor disagree
- ○ Somewhat agree
- ○ Strongly agree

Question 9. How do you evaluate this version?

Please choose one answer option for each of the categories.

	1	2	3	4		
lame	○	○	○	○	○	exciting
easy	○	○	○	○	○	challenging
complicated	○	○	○	○	○	simple
confusing	○	○	○	○	○	clear
not fun	○	○	○	○	○	fun
ugly	○	○	○	○	○	beautiful
ineffective	○	○	○	○	○	effective

Question 10. What is it that you LIKE about this version?

Please answer with at least one complete sentence. The sentence should be logically and grammatically correct and relate to the image above.

_____

_____

Question 11. What is it that you DO NOT LIKE about this version?

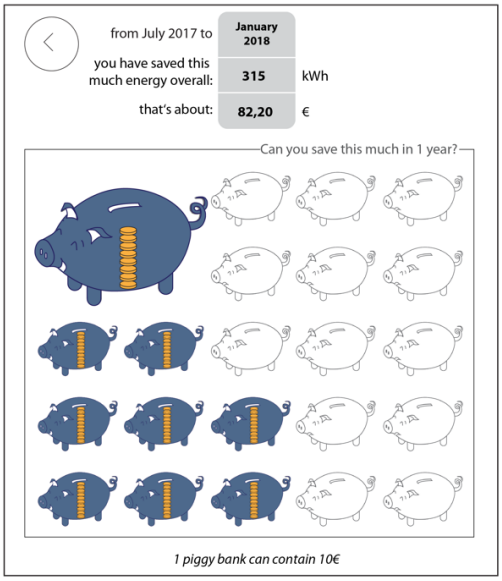
Please answer with at least one complete sentence. The sentence should be logically and grammatically correct, relate to the image above and differ from the sentence you wrote in the previous question.

_____

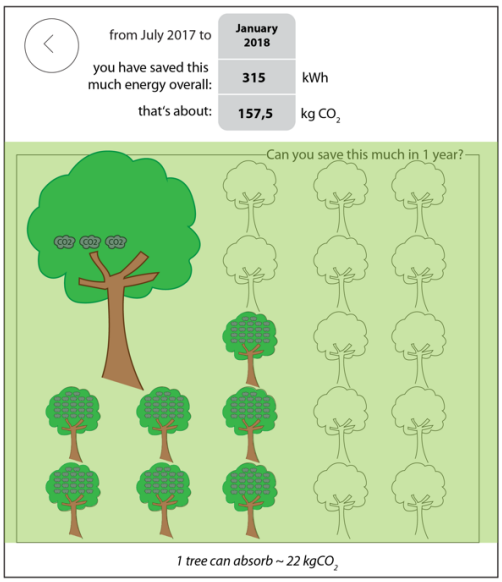
_____

PAGE 5^. Imagine you had to choose which of the versions you would like to be displayed by default in your energy saving application.

Visualization 1



Visualization 2



Question 12^. If you think about the visual design, which version do you prefer?
Please choose one answer option.

- ○ Strongly prefer 1
- ○ Moderately prefer 1
- ○ Neutral
- ○ Moderately prefer 2
- ○ Strongly prefer 2

Question 13^. Please explain the choice you made in the previous question.
Please write at least 2 full sentences. The sentences should be logically and grammatically correct and relate to the screenshots.

_____

_____

Question 14^. If you think about how <u>important</u> saving money and protecting the environment are to you <u>personally</u>, which version do you prefer?

Please choose the answer that best matches your opinion. Please notice that this question asks you to take a different standpoint for your decision.

- ○ Strongly prefer 1
- ○ Moderately prefer 1
- ○ Neutral
- ○ Moderately prefer 2
- ○ Strongly prefer 2

Question 15^. Please explain the choice you made in the previous question.

Please write at least 2 full sentences. The sentences should be logically and grammatically correct, relate to the screenshots and differ from the explanation that you gave above.

_____

_____

Demographic Questions. (Please refer to the question in Appendix 2 – they were the same as in the Overview Visualization test).