

Evaluating the Effect of Gamification on the Deployment of Digital Cultural Probes for Children

I. Rodriguez, A. Puig, D. Tellols, K. Samsó

*Departament de Matemàtiques i Informàtica.
Facultat de Matemàtiques i Informàtica. Universitat de Barcelona.
UBICS and IMUB research centers.
Gran via de les corts catalanes, 585, Barcelona, Spain*

Abstract

This research proposes a digital Cultural Probe (CP) for children in the context of energy use and awareness. A CP is a user research method that allows participants to self-report by means of tools, artefacts, and tasks that they complete at their own pace. Our CP aims to gather data about the energy related habits and knowledge of the children and their families and also aims to provoke reflection and raise their awareness of energy and environmental issues. It is well known that the main drawback of CPs is their low completion rate and the consequent lack of gathered data. Based on the impact of gamification on users' motivation and engagement, we aim to test the hypothesis that "Gamification may encourage an active and fruitful participation of children in Cultural Probes". CP consists of four tasks designed to gather information about the children and their families' electric consumption and saving habits. We developed the CP as an Android app which was evaluated with children from two primary schools in Barcelona, who interacted with two different app versions (gamified and non-gamified). We designed a comparative experiment that gathered data on CP completion rate, CP task completion rate, and the time required to complete the CP. To our surprise, we found a slight difference in favour of the non-gamified app in all of these measures, though this difference was not statistically significant. However, the difference in the children's awareness of and attitude to energy consumption before and after the realisation of the CP was significant as measured by the Wilcoxon signed-rank test. This was not the case for the effect on energy awareness of the children's relatives. The lessons we learned, which may also prove to be useful for researchers that aim to design cultural probes, are that user research methods for children can be effective enough when designed to be interactive and dynamic (i.e. digital CPs) and that the number of CP tasks may also be a factor to take into account. When the number of tasks is low, the CP is likely to achieve a high completion

*Corresponding author

Email address: inmarodriguez@ub.edu (I. Rodriguez, A. Puig, D. Tellols, K. Samsó)

rate and the additional effort of gamifying the application may be not worthwhile. Then, gamification may support a CP with a large number of tedious tasks. When designers want children to reflect on key issues, the CP should be thoughtfully designed because it is not easy to foster children’s reflection and reasoning.

Keywords: Cultural Probes, Gamification, User Research

1. Introduction

Cultural Probes (CP) gather data about people’s feelings, thoughts, interactions and knowledge, in order to inspire and inform the design of products [1]. They are a well-known method for conducting contextual research in the HCI community [2][3]. The literature includes works using CPs as they were initially conceived, that is, for design inspiration (with playful, open-ended, evocative and playful tasks) [4] [5] [6], focusing heavily on gathering qualitative data. Other works adapted this initial approach and used CPs as a means of collecting information (with more focused tasks, gathering mainly quantitative data), complementing data gathered through other methods such as observations and interviews [7] [8] [9].

Designed probes covered a variety of topics of interest, for example the so-called domestic probes [10], which were used to inspire the design by exploring people’s home lives in an intimate and sensitive context, and technology probes [11], which were employed to explore the use of low-fi prototypes and open up new design spaces.

Furthermore, although CPs were originally designed as a physical package including artefacts (cameras and diaries to record feelings and events) and tasks to be performed by users at their own pace, CPs can also take the form of digital applications, i.e. digital Cultural Probes [12].

Nevertheless, CPs upset users’ routines, with the risk of presenting a low completion rate [13], and varying the return rates of materials in different settings and populations, as reported by Gaver et al. [1]. Indeed, participants in a study on migrants in Australia gave reasons for non-completion such as a lack of time or a lack of interest [14].

In this research, we advocate for gamification - the use of game elements in non-game contexts [15] - as an additional component of playfulness in CPs to improve their completion rate. More concretely, we propose the following hypothesis: "Gamification may encourage a more active and fruitful participation in Cultural Probes". We tested this hypothesis by developing a digital CP related to energy use in families, which is integrated in a project of collective energy awareness. The goal of the project is to involve users actively in the Smartgrid, providing them - as prosumers, who buy and sell energy - with tools and solutions for energy management. More specifically, our digital CP aims to explore a family’s daily routines to design an application that supports them in an efficient use of energy.

However, the design of the applications related to resources management (i.e. energy, water) requires a sound knowledge of users' needs, values and habits. Additionally, in the sensitive context in which they are used (i.e.. households), other user-research techniques, such as interviews and direct observation, can be tedious to perform as the members of the family have different schedules and may be perceived as being somewhat intrusive. In this case, the use of CPs is an adequate user research technique for obtaining relevant information about families and their energy behaviour.

We decided to target families' children rather than adults for several reasons. First, children are naturally creative and this fact would foster the potential for surprise and in consequence the potential of uncovering unexpected findings [16]. Second, children usually have more free time and are in general more interested in this kind of hands-on activities than adults. Third, the direct participation of children in the CP - being in charge of gathering information about energy from other members of the family - can raise their awareness on the importance of a responsible use of energy at home, and can influence their families to adopt energy efficiency measures, as demonstrated by previous researches [17] [18]. We then provided primary school children (between 11 and 12 years of age) with both a gamified and a non-gamified version of the CP, and performed a comparative analysis of the results obtained.

In previous works, we proposed a draft of the CP tasks [19] and the gamification designs, including digital CP mock-ups [20]. In this paper, we further detail both designs (CP and gamification) and present the deployment and the evaluation of gamified and non-gamified versions of the CP. Our goal is to assess the effect of gamification on CP and CP's tasks completion rates, and on the time it takes to complete the CP. Additionally, we want to know whether performing CP tasks raises the energy awareness of the children and their families. Specifically, the use of a narrative and role-play in the gamified version of the CP aims to engage the participant in a mission to fight an evil wizard, in which they get rewards every time they complete a task/mission and, to add even more motivation, they compete with other children to win a final prize (being able to see the other children's scores on a leaderboard).

Our main focus of interest is how gamification can lead the children to obtain better CP results in terms of completion rate and performance time, while the gamification of a CP with the aim of inducing an energy behaviour change, as other researches have done [17] [21], would fall outside the scope of our research. In our case, it is the design of CP tasks (whether the CP is gamified or not gamified) that aims to collaterally raise awareness in the children and their families. From our results, we provide answers for those HCI researchers who raise questions regarding the CP design process such as *"To what extent is gamification useful in the design of a CP?"*, *"Would gamification foster CP participation in terms of completion rate and time to complete the CP?"*, *"Can CP design have an impact on participants' energy awareness?"* Specifically, we found slight differences between gamified and non-gamified CPs but they were not statistically significant. Moreover, our results showed that the digital CPs had an impact on the children's energy awareness though not on that of their

families.

This paper is structured as follows. Section 2 presents the state of the art, Sections 3 and 4 give an overview of CP tasks and gamification designs respectively, Section 5 presents the evaluation of the CP and discusses the results obtained, and the Conclusions section summarises our major findings and suggests lines of future research.

2. RELATED WORK

Gamification for behaviour change has attracted the interest of the research community in recent years. We can find studies in several domains, including health [22], energy [23], business [24], and education [25]. Even in the field of experimental design [26], a gamification study showed an increased completion rate of experimental tasks with children from 73% to 97%. Whatever the context, the goal is to engage users in fun and motivating activities that shape their habits and influence their behaviour [27] [15]. In this research we use gamification to enhance users' involvement and participation in cultural probes. Since our case study relates to energy use and consumption in families, we first review research works that proposed probes targeting children and families. We then present gamification studies in the energy domain.

2.1. Cultural probes with children and families

Since their emergence, CPs [1] have been used in several types of projects related to the domestic context. In this context, the probes design faced challenges such as the heterogeneity of roles in families ([9], [28]) and the disruption of the domestic routines ([11], [29]).

Vetere et al. [9] designed the magic box, a cultural probe to explore playfulness between children and grandparents when they are separated by distance. Researchers found several artefacts used to mediate play and the different roles adopted by both younger and older participants, recommending the design for its flexibility, since it allows for the expression of the different roles. Moreover, Oumard et al. [28] studied the need to stay in touch in families. They designed a package for heterogeneous individuals (young children, adults, teenagers) with different ages and motivations. They designed a package where each day involved a theme, with serious content but a playful appearance. The data gathered was diverse reflecting different viewpoints of the members of the family. The probes served the purpose of empathising with the entire family. Nevertheless the probes were difficult to analyse and they only adapted one concept in their project based on their findings.

Shewbridge et al. [11] explored the potential uses of 3D printers at home. They asked users to locate a fake 3D printer in a visible location at home, and gave them pens, index cards, and cameras. The results showed desirable uses of 3D printers such as replicating existing objects, modifying and customizing existing objects, and creating new custom objects. Another cultural probes study [29] investigated the role that Internet video plays in online social life and

media consumption at home. It included tasks such as taking photos (objects and places and where to spend time using devices) and writing diaries (about their daily routine and use of media use). They concluded that computers are users' companions of users during the most of time they are at home, and that Internet video helps them to socialise and is integrated in the daily life of heavy Internet users.

In relation to research works using CPs with children, the studies have followed the classical (physical package) approach, as well as the digital one. Wyeth [13] proposed a probes consisting of an activity pack (with instructions to draw the future classroom, design a day's outing and rate subject areas, among others) which was distributed to children in the school. This pack aimed to provide insights on children' mental models of classrooms, science learning, gadgets to assist learning and even future career aspirations. Another study [8] designed a cultural probes for children to gain insight into how children experience technology. The study used animal stickers that children had to relate to technologies. Nevertheless, as reported by the authors, the study failed at the abstraction level with the stickers, which seemed too difficult for children. Iversen et al. [12] designed a digital probe involving the use of mobile phones for reporting children's educational experience. Their results underline the benefits of using mobile devices in CPs, which spontaneously motivate children, since technology is strongly integrated in their daily life. Authors also highlighted the possibilities of the automatic tracking of the probe, as well as data collection and analysis.

What differentiates our CP from other contributions in the state of the art so far can be summarized in the questions: "Who is the package for?" and "Whose information is this?" CPs usually targeted people who reveal insights about themselves. Our proposal involves giving the CP to children, who are responsible for performing the indicated tasks and through them we obtain information about the energy consumption habits of the children and their families' energy consumption habits and their knowledge of and interest in energy related issues.

To the best of our knowledge, few research works have used a game-based design in CPs. Bernhaupt [30] focused on the study of new forms of media usage at home. The results were positive in some aspects: the game helped participants to focus on the topic and the participants who played a card game provided more answers to researchers' questions than those using the non-playful format. Nevertheless, the game influenced neither the number of photos or videos taken by participants nor the number of materials used by them, perhaps because of its simple hedonic design instead of a thoughtful gamified design. Another investigation [31] studied the relevance of tangible gamified probes to fostering sociability and a sense of progression and control in a cooperative learning process in the classroom.

Our research differs from prior works because they did not use gamification but card-based games in probes, which are essentially different approaches within the broader field of hedonic design. Gamification includes game elements in a non-game context (i.e. a cultural probe), while a game establishes rules and a goal in an entertainment context. Moreover, they designed physical CPs

while we focus on the study of digital gamified CPs.

2.2. Gamification in the energy domain

With the deregulation of energy and the emergence of new ways of producing energy at home numerous gamified and game-based applications have appeared with the aim of raising awareness of energy usages and needs [32]. Usually, these gamification proposals address information gaps on energy concepts, support learning, and aim to motivate behaviour changes in different environments (workplace, schools and households).

Gamified experiences started in large companies with the aim of encouraging sustainable behaviours among employees using cards, points, badges and leaderboards in collaborative and competitive activities [33] [32]. Currently, governments are supporting energy initiatives in schools through gamified learning activities. For example, the European project GAIA proposes the use of the IoT (Internet of Things) and experimental activities to teach about energy. They provide students with a lab kit activity - covering aspects of energy consumption and efficiency inside school buildings - together with a serious game. The preliminary findings are promising but further evaluation is needed to assess the contribution of the gamification. The KYEN (Know Your Energy Numbers) program [18] also teach young tweens to use energy tools to measure their family's consumption, together with general energy-related concepts. Their results highlighted the necessity of collecting behavioural data about energy use in order to improve learning. Undoubtedly, we think that probes would be a suitable method to gather these data.

The emerging use of smart-meters in residential houses has also led to the development of novel technical architectures which consider several game elements (e.g. rewards system, social connection, performance status) [34], as well as mobile game-based apps such as PowerAdvisor [35], PowerExplorer [36], PowerAgent [37] and Social Power [38], among others. The game's challenges are related to housework, such as cleaning or cooking, asking players to compare their own consumption to that of other players. In addition, several web-based gamified applications, which foster energy savings by sending reports from home, have evaluated their short-term and long-term effects, concluding that the resulting increase in awareness is temporary and tends to decrease after a short time [39], [21]. Actually, even though gamification in the field of domestic energy consumption seems to be effective in inciting people to save energy, the evidences provided by most of the gamification approaches are more indicative than conclusive, i.e.. experiments have a limited number of samples and low quality of evidences [23], [40].

Moreover, a recent analysis of 57 energy-related apps designed for young people reveals that the number of gamification and game design elements correlates strongly to the app ratings, and indirectly to app usages [41]. They considered 10 gamification elements in terms of motivational affordances [42] - points, leaderboards, badges/achievements, levels, stories/themes, clear goals, feedback, rewards, progress, and challenge - concluding that most of the apps

underutilise gamification components, using in average one gamification element. In general, designed platforms use gamification in a thin layer (using PBL - Points, Badges and Leaderboards). Thus, it is not easy for designers to reach the emotional side of users and their habits. In this research we aspire to achieve deep gamification, i.e. covering additional game elements such as the aforementioned: stories/themes, feedback, rewards, progress and challenges [43].

As a final remark, we note that previous energy-related projects and applications are focused on energy literacy and on influencing people's behaviour [44], instead of eliciting and inspiring the design of energy-related applications, as we do with our cultural probe. Even though our objective is different we were somewhat inspired by them since all of us are using gamification in the energy context.

3. CP DESIGN

Although the initial drafts of the probes [19] were intended for a wider range of ages, children between 8 and 12 years old, we opted for targeting slightly older children in the final design (with ages between 10 and 12 years). One reason for that is that the children had already learned at school what energy saving means, they knew that electricity costs money and knew that there are various sources of energy. Moreover, we wanted to ask them for reflections, and believed that children of these ages would be better at expressing and writing their thoughts. Nevertheless, we were aware of the fact that, while performing the CP, they could still have little idea of some issues related to energy use at home such as the frequency of use of some appliances. We therefore encouraged children to interact with their parents, asking them what they did not know. Additionally, where appropriate, the design of the probes should differentiate whether the children were interacting either with adults in the family or with their siblings (more details in Section 3.3). For example, the behaviour "long pre-heating of the oven" was included in the list of possible behaviours for adults and not in those for children.

CP design was conducted by an interdisciplinary team of energy experts, who offered their advice on energy related content, and HCI experts. The aim of the CP was to trigger inspiration and then gather relevant data from children and encourage a responsible attitude to and reflection on energy use and consumption. To achieve the CP goals, we defined 4 tasks through which children could be made aware of their family energy habits. Moreover, to better immerse kids in the subject of energy saving and green behaviour, the visual design of the digital CP was based on a childlike visual theme using green tones and minimalist fonts (see Figure 1).

Table 1 summarises the design of the CP, including a brief description of the tasks, the elements needed to complete them, and the rationale behind the tasks (from our perspective as designers) [19].

The CP is structured in four successive phases: introduction, registration, performing the CP tasks, and CP final.

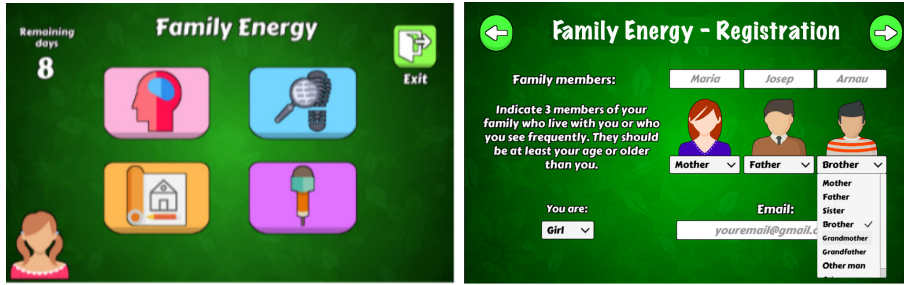


Figure 1: Digital CP: four Tasks and CP registration.

Table 1: CP tasks, the elements provided to perform them and their goals.

CP TASK	DESCRIPTION	ELEMENTS	RATIONALE
Task 1	Ask family members which electrical appliances they use the most and how they feel when using them.	List of family members, list of electrical appliances, list of moods.	Obtain information on users' emotional states to influence habit changes. Get the children to reflect on their responses.
Task 2	Observe family members to indicate their energy related behaviours and how good they are.	List of family members, list of behaviours, Open question "Do you think that <family' member name's> behaviour is respectful with the environment?"	Find out users' energy related habits and behaviours, find out users' level of commitment to environment and saving, and raise awareness. Get children to reflect on their responses.
Task 3	Identify the electrical appliances in each room of the house and specify their usage frequency on working days and weekends.	List of rooms, list of electrical appliances, rank of frequencies, open question "Do you think that your use of electricity is different on weekdays and weekends? Yes/No why?"	Find out users' frequency of use of electrical appliances, obtain usage habits from families as a whole, raise awareness. Get children to reflect on their responses.
Task 4	Ask energy related questions to family members and record their answers.	List of family members, closed-ended questions as a whole, multiple choice answers.	Find out users' knowledge, needs of information and preferences in terms of energy use, identify users' predisposition to changes.

3.1. First phase: introduction of the CP

This phase started with a brief text explaining the aim of the CP (i.e.. "help us get to know your family's energy knowledge and habits") and asking the children to perform the tasks. We also informed the children of the maximum number of days they disposed of to complete all of the tasks.

3.2. Second phase: CP registration

During registration, players identified three members of their family whom they would interact with in the CP. These members could be relatives who lived with them or who they saw frequently. They assigned an avatar to each participant (see Figure 1).

The digital CP offered the possibility of adding a mother, father, sister, brother, grandmother, grandfather, and another male and another female. Each

of them was assigned a faceless representative icon with the sole purpose of helping players to identify the family member through the CP. Moreover, kids had to introduce their relatives' names, which were used to personalize messages and content during the CP experience.

Once registered, users were able to complete the CP tasks in any order. Overlapping between tasks was possible: more than one task could be performed at any time, giving them the freedom to explore the whole set of tasks and not be forced to tackle them in a strict sequential order.

3.3. Third phase: performing the CP tasks

Each task was identified with an icon¹ and a color - pink, blue, orange and purple - to help the children be constantly aware of where they were, and what they were doing. We then gave them the relevant details for each task. The dashboard, see Figure 1, showed the countdown of days and the icons for accessing CP tasks.

- In task 1, the children gathered information about the electrical appliances that they and their relatives use the most and how they felt when using them. Concretely, they identified five combinations (appliance – mood) for each family member (see Figure 2). Based on the OCC model of emotions [45] which proposed a wide range of emotions, we selected basic ones (happiness, anger, and sadness) to facilitate the children's selection of feelings and interactions. We also included boredom, as proposed by [46], to cover situations that do not fit with basics emotions such as, "A father feeling boredom when using the vacuum cleaner".

Through this task, we obtained emotional information about the participants and their family to guide us when deciding how to influence their change of habits. Tips (or the way they are given) to save energy with an appliance can be more or less effective depending on the mood a person is in while using that appliance, For example, influencing a change in the use of some appliances can be easier if people feel negative about them and it may be necessary to explore other less or non-electrical alternatives to the use of those appliances that they feel positive about. Moreover, to foster reflection on feelings, once the children had finished the task, they could answer different open-questions, such as, "Why do you think the person feels that way"? The children could then either give their own response or ask about and discuss the issue with their relatives.

- In Task 2 children reported the energy-related behaviours of each family member. Each behaviour was described as positive (good behaviour) or negative (bad behaviour). An example of the positive use of devices is "Turns off the TV, computer or console when it is not being used", and its

¹Icons used to enhance the visual aspect of the app were obtained from Flaticon. <https://www.flaticon.com/>

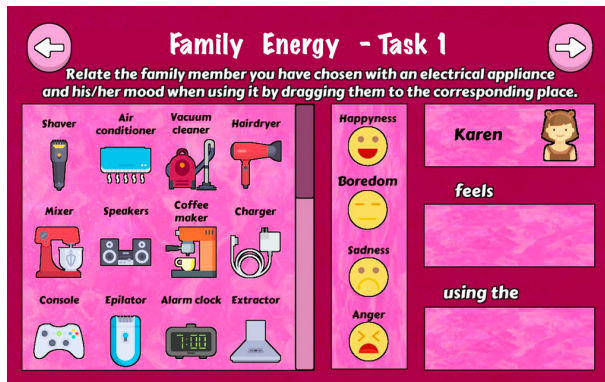


Figure 2: Task 1 about technical appliances that users use the most

negative counterpart is "TV, computer or console on standby when it is not being used". There were different messages for adults and children. Some good behaviours for adults were defined as follows: Some good behaviours for adults are defined as follows:

- Starts a full laundry load or doesn't do the laundry.
- Lets clothes dry outdoors or doesn't dry them.
- Preheats the oven for the indicated necessary time or doesn't use it.

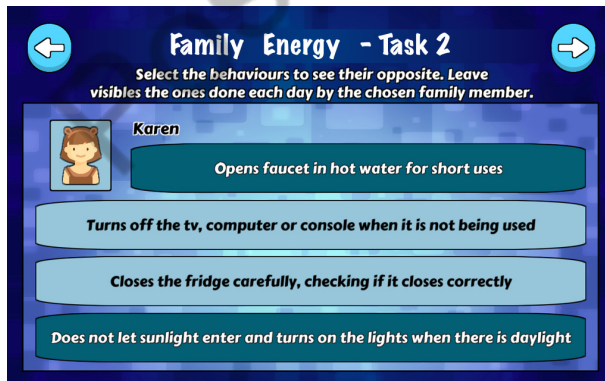


Figure 3: Task 2 to indicate energy-related behaviours.

Then, kids could touch the positive description and change it to the contrary message (bad behaviours):

- Starts with a non-full laundry load.
- Dries clothes with a dryer.

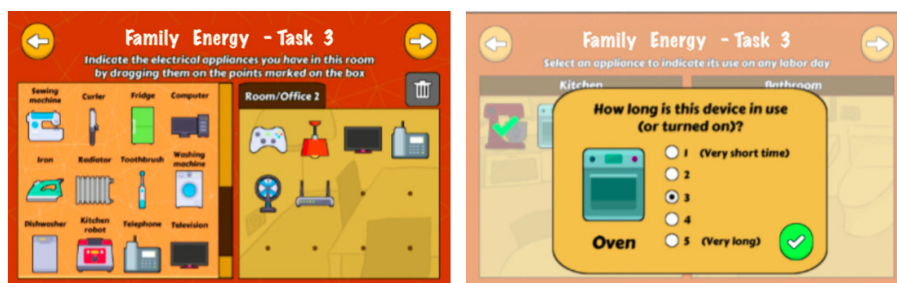


Figure 4: Two parts of task 3 where kids locate electrical appliances in their home.

- *Preheats the oven for longer than necessary.*

On the other hand, examples of good behaviours for children are:

- *Opens faucet in cold water for short uses.*
- *Turns off the TV, computer or console when it is not being used.*
- *Ventilates the house a few minutes every day opening windows or balconies.*

And examples of the corresponding bad behaviours include:

- *Opens faucet in hot water for short uses.*
- *Turn on the TV, computer or console when it is not being used.*
- *Doesn't open windows or balconies to ventilate the house a few minutes a day.*

Once behaviours had been selected from the list, the children were asked to reflect on the behaviour of their family member. First, they answered a closed question "Do you think <family-member-name> has a responsible behaviour?" Yes/No, and after, answering "Why do you think it is a responsible/irresponsible behaviour?".

- In Task 3 the children identified the electrical appliances in each room in their house (see top of Figure 4). After that, from the previously added appliances, they had to indicate their frequency of use on working days (any day from Monday to Friday on average) and weekend days (any Saturday, Sunday or public holiday) (see bottom of Figure 4).

Afterwards, to foster their reflection, a panel showed up asking them the following questions: "Do you think that the use of electricity is different on working days than on weekend days at your house?" (closed question with options "Yes" or "No", and then an open question "Why?"); "In what room do you think the use of electricity is excessive and what would you do to reduce it?" (open question with an optional response).

- In Task 4 children interviewed each family member. As children are not skilled at conducting an unstructured interview, we designed a structured one with closed questions. We designed different multi-choice questions for adults and children.

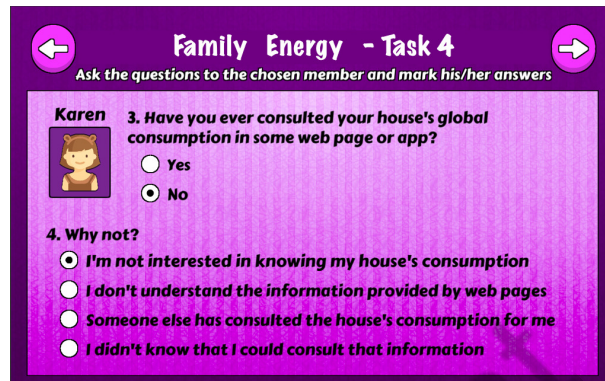


Figure 5: Task 4 where kids interview their relatives.

Some of the questions in the interview for children were:

- *Do you know how much money is spent on the household electrical bill?*
 - * *Yes.*
 - * *No.*
- *If saving energy at home had a positive impact on you as a reward, would you apply steps to reduce electricity consumption?*
 - * *No, I already do my best to save energy.*
 - * *Yes, I think I can try harder to save energy.*
 - * *It depends on the kind of reward.*

The corresponding questions in the interview for adults were:

- *Have you ever consulted your house's global consumption in some web page or app?*
 - * *Yes.*
 - * *No.*
- *(If "Yes") What web?*
 - * *The energy company's web/app (CompanyX, CompanyY, etc).*
 - * *A general web/app of consumption calculation.*
 - * *A web/app connected to the house's meter that allows me to see my consumption in real time.*

Finally, the CP ended when all the tasks were completed or when the time was exhausted. The children had eight days to complete the entire CP.

This CP design was implemented using the Unity ² game engine, a development platform for building 2D and 3D games. Concretely, our development used Unity 2D API and followed a client-server architecture to store children's progress and the user research data gathered during CP deployment in a remote database system.

4. Gamified CP design

The aforementioned CP tasks corresponded to four missions in our gamified CP design. Participants immersed themselves in a role-play and story (setting and characters) and obtained rewards for completing the missions. The children played the part of special agents who fought against an evil wizard in four missions in which they were disguised as a psychologist, a detective, an electrician and a journalist.

Each mission had a limited time for its completion, and the rewards for players were KWh (i.e. a measure of energy) and badges. Moreover, the children had to accept missions in order to begin the countdown. Once the mission had been accepted, it had to be completed within two days to obtain the maximum reward and the corresponding badge. Otherwise, players only receive the 60% of the points assigned to that mission.

In the following we provide details of the aforementioned CP phases for the gamified CP.

4.1. *First phase: introduction of the CP*

The gamified CP starts with an introduction to the story "*You are a special agent selected from the Energy Protectors' Agency to become part of a very important operation to save people wasting energy like crazy, enchanted by an evil spell.*" The challenge starts when the evil wizard steals 1.000.000 KWh which the players (with 0 KWh) have to recover in their role as a special agent.

4.2. *Second phase: registration*

Kids register in the gamified CP as previously detailed in section 3.2, i.e. identification of relatives and avatar selection.

4.3. *Third phase: performing of CP missions*

In the app dashboard, we added gamified elements (highlighted in yellow in Figure 6): the child's avatar, their progress (points, badges and a leaderboard), and the current status of each one of the missions. From here they could also consult a brief text setting out the rules, change their avatar, and access the leaderboard to see where they ranked compared to their classmates. Figure 7



Figure 6: Dashboard of gamified CP.

shows the design of the leaderboard, the points and the badges in the gamified app.

Furthermore, several icons appear on top of the mission buttons in the dashboard of the gamified app:

Lock: This means the mission has been accepted, so it is an indication for the children that the mission has started.

Clock: This means the mission has been accepted and there is still time to complete it to get the full reward points and the corresponding badge. This icon is also a button that allows players to see how much time is left to complete the mission in time.

Check: This means the mission has been completed.

Badge: This means that the mission has been completed in time and the player has won the corresponding mission badge.

The specific mechanics of the missions are set out below:

- To complete the psychologist mission children had to specify 5 combinations (appliance – mood) for each family member, for which they obtained 40.000 KWh (points) and, if this task was completed in time, they received a badge consisting of a psychologist’s diploma. The whole mission was worth 120.000 KWh.
- To complete the detective mission, the children had to identify the behaviours of each family member, for which they obtained 60.000 KWh and, if the task was completed in time, they won a pair of detective’s binoculars as a badge. The whole mission was worth 240.000 KWh. Finally, players receive feedback on that family member’s behaviour (quite responsible, not very responsible or balanced) based on how many good or bad behaviours they had selected in the task.

²<https://unity3d.com/>



Figure 7: Points (KWh), badges and leaderboard screens

- In the electrician mission, children received 80.000 KWh when they identified the electrical appliances in each room in their house (see top of Figure 4). After that, from the previously added appliances, they have to indicate their frequency of use on working days (any day from Monday to Friday on average) and weekend days (any Saturday, Sunday or public holiday). For each type of day completed they won an additional 80.000 KWh. If the mission was completed in time, they earned a badge picturing some electrician's tools. The entire mission was also worth 240.000 KWh.
- To complete the journalist mission, the children interviewed each family member. They were rewarded with 90.000 KWh for each member they interviewed and, if is the task was completed in time, they were awarded a press accreditation as a badge. The entire mission was worth 360.000 KWh.

As previously introduced in CP tasks' description (Table 1), the first three missions included the gathering of children' opinions and reflections on their and their families' behaviours and habits.

Finally, players who recovered the most KWh from the wizard were the CP winners. If the player completed all their missions in time (thereby winning the 4 mission badges) they also obtained a special agent badge. Moreover, if the whole CP was fully completed in less than two days, the players recovered an extra 50.000 KWh. This was a surprise element that aimed to reward the dedication and involvement of those who showed more interest and willingness to participate by placing them at the top of the ranking.

5. EVALUATION

We performed a comparative study of the gamified CP and the non gamified CP. For this study, we collected quantitative data such as the completion rate, both at the higher level of the entire CP and at the finer-tuned level of each task/mission to be completed, and the time spent on each task/mission. We were also interested in analysing qualitative data and gathering children’s opinions on their experience, such as how much they had enjoyed the CP activities and their perception before and after performing the CP of their own and their family’s level of awareness of energy related issues.

5.1. Methodology

We performed a between-subjects test on two groups of students between 11 and 12 years of age, the experimental group (gamified app) and the control group (non gamified app). A consent form was signed by their parents of children who participated, who which gave information about data anonymity and the use of data only for research purposes³.

The collaboration of Rubi Brilla (energetic initiative of the town hall of Rubi) was essential for this part of the project. Thanks to them, volunteer children from two different schools were recruited to participate in the Cultural Probe evaluation. The children owned a mobile device and, since many of CP tasks required the interaction of the participants with members of their family, it was preferable that they lived in a house with three more relatives who were older than them. We recruited children from two different schools, since having both control and experimental groups in the same school may have distorted results, since the children in the non-gamified CP group would have been aware of other classmates using a gamified app. We think that this playful version of the app would introduce a bias in the CP completion rate in the non-gamified condition. Both schools were located in the same neighborhood, with the same social and cultural context and the same involvement in energy education.

Each app version (gamified and non-gamified) was randomly assigned to each school: Gamified CP - School 1 (14 participants), and Non Gamified CP - School 2 (11 participants).

The next step was to present the project at each of the schools, explain the basic concepts, briefly describe the purpose of the app and tell participants that the one who completed the CP the best (in quantity and quality) would receive a prize. By quality we meant that the children would write meaningful messages whenever asked for opinions and reflections during the CP. The existence of two different versions of the app was never mentioned in order to avoid conditioning the outcomes. The particularities (goals and tasks) of the CP were then described in the same way at both schools. The plan in this first session was to get the app installed on all the children’s devices and to ask them to fill out a pre-test questionnaire.

³Approved by University ethics committee.

During the eight-day long CP, the participants had to interact with the app in their homes while all the data was stored on a remote database server. At the same time, a daily tracking was performed to ensure that everything was working correctly.

We visited the schools two weeks after the introduction to summarise the information gathered through the CP tasks and to deliver the prize to the winners. Three winners were chosen from each school based on their score in the ranking (only in the gamified version) and the quality of their responses to the open questions that appeared throughout the app. The same day, participants were asked to answer a post-test questionnaire. The questionnaire was the same for both schools since the information to be gathered was the same, including the level of energy awareness before and after the CP and the most liked tasks, among others.

The app was also tested on a third school (Escola del Mar in Barcelona), but the resulting data were not used to measure rate completion results since we recruited an insufficient number of children for the non gamified CP. However, the results of the pre and post questionnaires of the participants who tested the gamified app have been aggregated to the information from the other two schools (see section 5.3).

5.2. Comparative results

In order to test the main hypothesis of this research, which claims that gamified CPs may obtain better results than non-gamified ones, we measured CP completion rates, CP task completion rates, and the time taken to complete the whole CP. Task completion rate is expressed as a value from 0 to 4 (from no tasks completed to all tasks completed) and times are represented in hours.

5.2.1. The CP Completion Rate

The results for CP completion rate, which indicates the percentage of participants who completed the CP, are shown in Table 2. The N-1 Chi-squared non-parametric test for small samples gives us a p-value of 0.8434, and given that it is much greater than 0.05 (for a confidence interval of 95%), the difference in the completion rates for both CPs is not significant enough to allow us to reject the null hypothesis. We therefore conclude that there is a low, 15.65%, chance that the rates are different (in fact, the non-gamified CP seems to have achieved better results).

CP	Gamified	Non-gamified
#Participants	14	11
#Completed	11	9
Completion rate (%)	78,57	81,82

Table 2: Task completion statistical values, both gamified and non-gamified

5.2.2. CP Tasks Completion Rate

To calculate the task completion rate, we recorded the number of tasks completed by participants. The rate is calculated in relation to the four tasks in the CP. To test the significance of the difference in the results obtained (Table 3), we performed an unpaired t-test (p-value=0,2473, ES=0,5 which is a small-medium effect size according to Cohen's d) [47]. As we can see, the difference in completed tasks between both CPs also fails to reach statistical significance.

TASK COMPLETION	mean	min	max	sd	rate (%)
Gamified	3,54	1	4	1,01	88,5
Non gamified	3,91	3,25	4	0,23	97,75

Table 3: Task completion statistical values, both gamified and non-gamified

5.2.3. CP Completion Time Rate

To calculate CP completion time, we measured the number of hours that participants spent using the CP until completion (see Table 4). Rate is calculated in proportion to the total of 192 hours that participants had to finish the CP (8 days).

In this case, the average results are similar, highlighting the fact that both versions of the CP were completed, on average, in less than half the time given. The statistical values used to perform the test are depicted in Table 4. The p-value obtained is 0,7335, indicating that the difference in the spent time to complete the CP in both versions is not statistically significant (taking a significance level of 0,05). Once more, the null hypothesis cannot be rejected.

CP COMPLETION TIME	mean	min	max	sd	rate (%)
Gamified	84,71	16,8	173,8	64,17	44,12
Non gamified	93,64	33,9	167,3	47,84	48,77

Table 4: Task completion time statistical values, both gamified and non-gamified

In conclusion, the null hypothesis cannot be rejected, but the alternative hypothesis cannot be accepted either. Therefore, further research should shed more light on the effect of gamification in CP results, for example by adding a new experimental condition (physical CP), by re-designing CP (new tasks), and, of course, by considering a larger sample size.

The following two subsections show the results of the pre-test and post-test questionnaires related to energy awareness and the data specific to the app, respectively. Note that the percentages given are for a total of 27 children (including those at the third school) who tested the app and answered both questionnaires.

5.3. Questions about energy awareness (Pre-test and Post-test)

We aimed to discover whether children’s awareness of energy related issues, such as saving and being responsible towards the environment, was higher or not after their experience with the app.

In order to find out whether performing the CP tasks motivated the children to discover more about electricity, their energy consumption and other ways to be respectful with the environment, we asked participants a number of questions in pre and post (two weeks after the end of the CP) questionnaires:

”Q1: What is your level of interest about knowing how much energy you consume and how much you are capable to save?”

”Q2: What is your level of compromise with a responsible use of energy?”

”Q3: What is your family level of compromise with a responsible use of energy?”

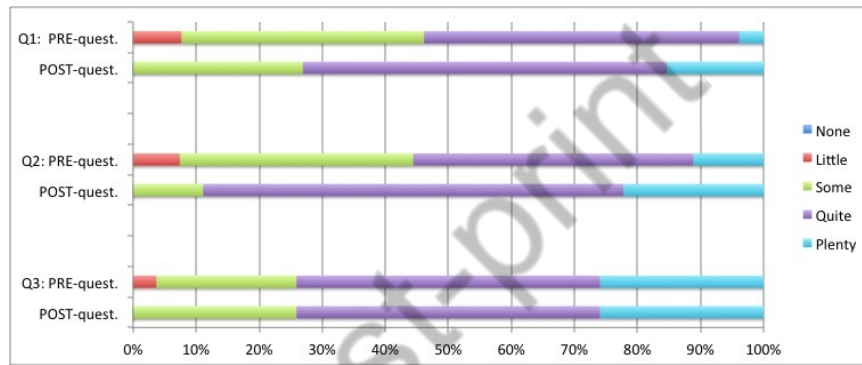


Figure 8: Pre-post test: 3 likert-type questions related to energy awareness

Figure 8 summarises the results for Q1, Q2 and Q3. We can see that the activities in the CP had an impact on children, raising their awareness and encouraging them be more responsible in their energy consumption. Nevertheless, it seems that the CP tasks did not affect the children’s relatives (at least from the children’s point of view). We analysed the data using the Wilcoxon Signed-Rank Test for dependent samples. This test gave results of $W = 8$ and $W = 5$ for Q1 and Q2, which are less or equal to the critical values of 10 and 15 respectively, for a one-tail test at a significance level of 0.05. There is therefore sufficient evidence to claim that the population median of differences is other than 0 and we can reject the null hypothesis. However, the Wilcoxon test showed that the children’s perception of their family’s awareness of energy was not significantly affected by their interaction with the app, with a value of $W = 42$, which is greater than the critical value of 17 ($\alpha=0.05$).

5.4. Questions about the experience

The questionnaire also included a small number of more specific questions about their experience. We categorized the children’s answers to ”What moti-

vated you to do the tasks of the app?" in two different topics:

- Energy-related: They wanted to be more careful with the environment, find out about their electric consumption and how to reduce it, and were interested in energy related issues.
- App-related: They had fun with the app (which would not have been the case if the CP had been given to them in its non digital version), they wanted to win a lot of points in order to be first on the ranking and win the final prize (only in the gamified version), and they were also curious to discover what exactly they had to do in each task (briefly explained in the presentation session performed at schools).

Moreover, in order to get more feedback, the children were asked about the tasks (which ones they liked and which ones they did not like) and they had to explain their experience when performing them, which could prove be useful for a redesign of the Cultural Probe.

Figure 9 (a) shows the result of the question "What task did you like the most?". The Electrician mission clearly stands out for being the most liked task. Below we summarise the reasons given by kids when selecting the most liked task:

- Task 1/Psychologist mission: They found it fun because they bonded with their family. They were not aware that using appliances aroused emotions on them.
- Task 2/Detective mission: They liked to investigate and discover how their relatives behaved.
- Task 3/Electrician mission: They liked being more aware of their electrical appliances and their usage. They found it to be the most entertaining task.
- Task 3/Journalist mission: They enjoyed asking their relatives questions.

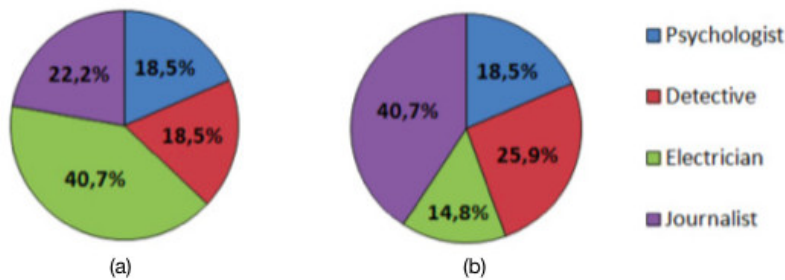


Figure 9: (a) Task they liked the most, and (b) Task they liked the less.

Figure 9 (b) shows the responses to the question *"What task did you like the least?"*. In this case, there is also one task that stands out for being the least liked: the Journalist task. Below, there is a compilation of the reasons given as to why they did not like the task selected.

- Psychologist task: They found it boring and/or absurd. We think that the reason for their boredom was that they had to repeat this task for each member of the family. Moreover, sometimes it was difficult to associate an emotion to a person using an electrical appliance. For example, the children did not know how they or their relatives felt when using a toothbrush.
- Detective task: They did not understand the mechanics of the task and failed to perform it correctly. Some of the participants told us that they did not know whether they had to "spy" on their relatives and observe their good and bad behaviours or whether they had to ask them. Indeed, we wanted the children to observe and try to recall their relatives' past actions but we also considered that asking them was also acceptable. Although the app had an option to ask for help, with an explanation of each task, they did not use it.
- Electrician task: They found it chaotic to go around the house looking for electrical appliances. We think that were several reasons why a number of children found the task chaotic, including the fact that they were tired, the dimensions of their houses and a lack of time. Nevertheless, this task was the one that the children enjoyed the most.
- Journalist task: They found it too long to be completed (too many questions) and they got tired. Their relatives did not have enough time to answer the questions. They generally do not like conducting interviews.

In regard to the question *"What didn't you like about the app?"*, some negative, but really useful, feedback was obtained. Nevertheless, 11 children indicated that they had liked everything about the app (many of them having access to the leaderboard to see their classmates' scores). The children's answers are classified in two different topics:

- CP design: They found it too short (only 4 activities). They claimed they would have liked to continue performing more tasks. They did not know how to answer to some of the questions asked through the CP.
- Gamification design: They did not like the fact that extra points were going to be awarded (this was a surprise element of the CP presented in the Mechanics paragraph of Section 4.2. Gamification). They did not like the narrative.

Finally, to try to understand the children better and learn from their ideas, they were asked: *"How do you think the app could be improved?"*. Although a

number of children claimed that they liked everything, they did point out some possible improvements to the app:

- Adding more tasks to make the app last longer.
- Changing (or reducing) the questions asked through the CP.
- Explaining specific energy-related terms that may be unfamiliar to them.
- Adding the possibility of winning virtual goods or buying things with virtual currency.

We also analysed the data gathered through the CP. In the first CP task, it is noteworthy that the "Joy" emotion was used a lot more than the other three emoticons to describe how the children and their relatives felt while using an appliance.

In the second task, it was surprising to observe that most of the participants and their relatives displayed green behaviour and that those who did not acknowledged their mistakes and wanted to change their habits.

In the third task, it was interesting to note that all the children answered that, indeed, there was a difference between their use of electricity on working days at the weekend, though from two different points of view. A number of children claimed that electricity usage was higher on working days because they spent the weekends going out with their families. On the other hand, the rest of the children stated that their electricity usage was higher at the weekend because they had the chance to play video-games, watch TV and cook, among other activities. These are things they could not do on working days because they spend most of the time at school or at work.

Lastly, in the fourth task, the answers revealed that most people did not know they could consult their electricity consumption or did it through their electric company. Participants thought they could reduce their consumption without much effort. They manifested that they understand consumption data better in Euros, and they agreed that their children would contribute to reducing their consumption if they obtained a reward in return.

5.5. Discussion

The results reported for the gamified and non-gamified app in terms of CP completion rate, CP task completion rate and time to complete fail to shed light on the supposed benefits of gamification. We believe that kids found the interactive nature of the CP (digital CP) and its close relation to their family's day-to-day life attractive. One child told us that, "it was fun to talk about energy with my relatives". Moreover, the low number of CP tasks may have influenced the high rate of completion in both experimental conditions [13]. Although we told teachers not to insist on the completion of CP tasks, a slightly (uncontrolled) difference in involvement by the schools may have led to a higher or lower completion rate. Finally, the number of participants was low.

Hamari's study concluded that the effects of gamification are greatly dependent on the users and on the context in which gamification is deployed [42].

Nevertheless, we are convinced that the deployment of CP in different schools did not influence our results since they were located in the same area and the children had similar sociocultural traits. In fact, the deployment of both experimental conditions in the same school may have produced distorted results since the children using the non-gamified app may have been influenced by the fact they knew that their classmates were interacting with a different, more "fun" version of the app. We believe that the CP completion rate in the experimental condition would have dropped radically.

Although we should take a bigger sample of individuals in order to draw firmer conclusions, our results suggest that gamification may not be so determinant when designing a CP for children, especially when they are digital and when the number of CP tasks is low. Additionally, gamification has proven to be effective in motivating kids to perform repetitive and tedious tasks [26]. Our recommendation is therefore to concentrate one's efforts on the design of the gamification, when there is the possibility of low completion rates and when the nature of the tasks is boring and monotonous.

We hope our work will stimulate further research in this area addressing questions such as whether digital gamified CPs for children yield better results than non-digital (physical) ones; how to design CPs for different target population (such as children and adults); the influence of the target population on CP results and the influence of the number of tasks and tasks design on CP results.

Regarding the influence of children on their families, previous studies reported the potential of changing parent energy-saving behaviours through child-focused interventions [17] [18]. However, our results suggest that children were more aware on energy related issues after the CP than other members of their family. One explanation for this fact can be found in the design of the CP, which did not actively engage other members of the family. The question then arises of what would happen if the design of the CP promoted the active participation of the children's relatives. Moreover, further research would be useful to evaluate the long-term impact on every member of the family [48].

Related to children's opinions on the most and least liked CP tasks, we think the electrician was the most liked task because it was the one that required the most input from the children, who had to select rooms and apparels and modify the map of the house, among other tasks. In contrast, they did not like the journalist task because it required them to read the questions and possible answers to their relatives. Indeed, studies agree that most children prefer dynamic and interactive tasks than those that are more cognitively intensive, involving reading and understanding [49]. Another issue raised during the evaluation was that mainly during the journalist task children sometimes did not know the meaning of a concept, which they found frustrating. As a result, we are designing a new version of the app that includes an embodied conversational agent, thereby allowing children to ask and talk about CP tasks and energy related concepts [50].

At the end of some of the tasks, the children were asked to reflect on the issue dealt with, and this did not prove to be popular. For example, we asked their opinion about how responsible their parents' habits were and why they thought

their relatives were happier using one device or another. We believe that these questions appeared suddenly during the interaction and that the children may have needed some help to encourage their reflection. In this context, we think that having a conversation with an embodied agent may motivate children and foster their reflections. Additionally, this help can be obtained from previous studies on sentence completion, a technique that produces qualitative data on users' views in a structured form [51]. As final remark, this research enforces the challenge of facilitating and fostering children's reflection and reasoning on resource management at home (energy, water, residuals).

6. CONCLUSIONS AND FUTURE WORK

The goal of this research is to understand the effect of gamification on Cultural Probes (CP) designed for children. A CP was developed with the aim of obtaining insights on the energy related habits and information needs of children and their families. The CP consisted of four tasks designed to gather information about their and their families' electric consumption and saving practices.

We designed an experiment to compare children of 11 and 12 years of age at two schools who were assigned to the experimental group (gamified app) and the control group (non-gamified app). We were unable to demonstrate the influence of gamification on the performance of CP tasks. Surprisingly, the non-gamified group obtained slightly better results than the gamified group, though this difference was not statistically significant.

As lessons learned we highlight the conclusion that user research methods for children are effective when they are designed to be interactive, visual and dynamic. The influence of additional elements to the CP, such as gamification, may have no clear effect on results. Designers should take into account the number of CP tasks because the additional effort of gamifying a CP when there are few CP tasks may not be worthwhile. We are also convinced of the influence of the nature of the tasks. The children in our study found the CP tasks to be exciting in both the non-gamified and the gamified formats, hence the good results obtained for both configurations. However, when CP tasks are perceived to be tedious, gamifying the CP experience may prove worthwhile. Finally, when designers want children to reflect on key issues, some extra help may be required to overcome the innate difficulty in fostering reflection and reasoning in children.

As on-going research we are designing an embodied conversational agent that accompanies children during the CP. This agent is designed to have structured conversations (i.e.. non-linear but branched dialogues) with children. Based on our previous work, this agent is endowed with personality, mood and emotions [50]. Concretely, the agent embodiment is an anthropomorphic earth. This time we aim to evaluate how effective and attractive the agent is for engaging children in conversations.

Acknowledgements Research supported by 2017-SGR-341, TIN2015-66863-C2-1-R, MIS-MIS-Language (PGC2018-096212-B-C33) and Accio COMRDI18-1-0010.

References

- [1] B. Gaver, T. Dunne, E. Pacenti, Design: cultural probes, *interactions* 6 (1) (1999) 21–29 (1999).
- [2] K. Boehner, J. Vertesi, P. Sengers, P. Dourish, How hci interprets the probes, in: *Proceedings of the SIGCHI conference on Human factors in computing systems*, ACM, 2007, pp. 1077–1086 (2007).
- [3] P. Wright, J. McCarthy, Empathy and experience in hci, in: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ACM, 2008, pp. 637–646 (2008).
- [4] F. Vetere, H. Davis, M. R. Gibbs, P. Francis, S. Howard, A magic box for understanding intergenerational play, in: *CHI'06 Extended Abstracts on Human Factors in Computing Systems*, ACM, 2006, pp. 1475–1480 (2006).
- [5] L. Kuiper-Hoyng, J. Beusmans, Using home networks to create atmospheres in the home: technology push or (latent) user need?, in: *Proceedings of the conference on Dutch directions in HCI*, ACM, 2004, p. 7 (2004).
- [6] D. Swallow, M. Blythe, P. Wright, Grounding experience: relating theory and method to evaluate the user experience of smartphones, in: *Proceedings of the 2005 annual conference on European association of cognitive ergonomics*, University of Athens, 2005, pp. 91–98 (2005).
- [7] A. K. Dey, E. de Guzman, From awareness to connectedness: the design and deployment of presence displays, in: *Proceedings of the SIGCHI conference on human factors in computing systems*, ACM, 2006, pp. 899–908 (2006).
- [8] K. Battarbee, A. Soronen, F. Mäyrä, Living in a zoo: bringing user experiences with technology to life, in: *Proceedings of the third Nordic conference on Human-computer interaction*, ACM, 2004, pp. 373–376 (2004).
- [9] F. Vetere, H. Davis, M. Gibbs, S. Howard, The magic box and collage: Responding to the challenge of distributed intergenerational play, *International Journal of Human-Computer Studies* 67 (2) (2009) 165–178 (2009).
- [10] F. Vetere, M. R. Gibbs, J. Kjeldskov, S. Howard, F. Mueller, S. Pedell, K. Mecoles, M. Bunyan, Mediating intimacy: designing technologies to support strong-tie relationships, in: *Proceedings of the SIGCHI conference on Human factors in computing systems*, ACM, 2005, pp. 471–480 (2005).
- [11] R. Shewbridge, A. Hurst, S. K. Kane, Everyday making: identifying future uses for 3d printing in the home, in: *Proceedings of the 2014 conference on Designing interactive systems*, ACM, 2014, pp. 815–824 (2014).
- [12] O. S. Iversen, C. Nielsen, Using digital cultural probes in design with children, in: *Proceedings of the Conference on Interaction Design and Children*, IDC'03, NY, USA, . ACM, 2003, pp. 154–54 (2003).

- [13] P. Wyeth, C. Diercke, Designing cultural probes for children, in: Proceedings of the 18th Australia conference on Computer-Human Interaction: Design: Activities, Artefacts and Environments, ACM, 2006, pp. 385–388 (2006).
- [14] S. K. Robertson, Cultural probes in transmigrant research: a case study, *InterActions: UCLA Journal of Education and Information Studies* 4 (2) (2008).
- [15] S. Deterding, Gamification: designing for motivation, *interactions* 19 (4) (2012) 14–17 (2012).
- [16] J. Hoffmann, S. Russ, Pretend play, creativity, and emotion regulation in children., *Psychology of Aesthetics, Creativity, and the Arts* 6 (2) (2012) 175 (2012).
- [17] H. Boudet, N. M. Ardoin, J. Flora, K. C. Armel, M. Desai, T. N. Robinson, Effects of a behaviour change intervention for girl scouts on child and parent energy-saving behaviours, *Nature Energy* 1 (8) (2016) 16091 (2016).
- [18] M. L. Mauriello, C. Zanooco, G. Stelmach, J. Flora, H. Boudet, R. Rajagopal, An energy lifestyles program for tweens: A pilot study, in: Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems, ACM, 2019, p. LBW1118 (2019).
- [19] D. Tellols, K. Samso, I. Rodriguez, A. Puig, Cultural probes for the gamification of energy awareness, in: *ACM WomENCourage*, 2016 (2016).
- [20] K. Samso, I. Rodríguez, A. Puig, D. Tellols, F. Escribano, S. Alloza, From cultural probes tasks to gamified virtual energy missions, in: Proceedings of the 31st British Computer Society Human Computer Interaction Conference, BCS Learning & Development Ltd., 2017, pp. 79–84 (2017).
- [21] D. Geelen, D. Keyson, S. Boess, H. Brezet, Exploring the use of a game to stimulate energy saving in households, *Journal of Design Research* 14 10 (1-2) (2012) 102–120 (2012).
- [22] P. Pereira, E. Duarte, F. Rebelo, P. Noriega, A review of gamification for health-related contexts, in: International conference of design, user experience, and usability, Springer, 2014, pp. 742–753 (2014).
- [23] D. Johnson, E. Horton, R. Mulcahy, M. Foth, Gamification and serious games within the domain of domestic energy consumption: A systematic review, *Renewable and Sustainable Energy Reviews* 73 (2017) 249–264 (2017).
- [24] D. Gears, K. Braun, Gamification in business: Designing motivating solutions to problem situations, *CHI* (2013) 4503–1952 (2013).

- [25] J. Majuri, J. Koivisto, J. Hamari, Gamification of education and learning: A review of empirical literature, in: Proceedings of the 2nd International GamiFIN Conference, GamiFIN 2018, CEUR-WS, 2018 (2018).
- [26] R. Brewer, L. Anthony, Q. Brown, G. Irwin, J. Nias, B. Tate, Using gamification to motivate children to complete empirical studies in lab environments, in: 12th International Conference on Interaction Design and Children, 2013, pp. 388–391 (2013).
- [27] K. Werbach, D. Hunter, For the win: How game thinking can revolutionize your business, Wharton Digital Press, 2012 (2012).
- [28] W. Horst, T. Bunt, S. Wensveen, L. Cherian, Designing probes for empathy with families, in: Proceedings of the conference on Dutch directions in HCI (Dutch HCI '04). ACM, NY, USA, 15-, 2004 (2004).
- [29] M. Oumard, D. Mirza, J. Kroy, K. Chorianopoulos, A cultural probes study on video sharing and social communication on the internet, in: Proceedings of the 3rd international conference on Digital Interactive Media in Entertainment and Arts, ACM, 2008, pp. 142–148 (2008).
- [30] R. Bernhaupt, A. Weiss, M. Obrist, M. Tscheligi, Playful probing: making probing more fun, in: IFIP Conference on Human-Computer Interaction, Springer, 2007, pp. 606–619 (2007).
- [31] R. Gennari, A. Melonio, S. Torello, Gamified probes for cooperative learning: a case study, *Multimedia Tools and Applications* 76 (4) (2017) 4925–4949 (2017).
- [32] F. Grossberg, M. Wolfson, S. Mazur-Stommen, K. Farley, S. Nadel, Gamified energy efficiency programs, American Council for an Energy-Efficient Economy, Washington, DC (2015).
- [33] K. Kuntz, R. Shukla, I. Bensch, How many points for that? a game-based approach to environmental sustainability, in: Proceedings of American Council for an Energy-Efficient Economy Summer Study on Energy Efficiency in Buildings, ACEEE, 2012, pp. 126–137 (2012).
- [34] T. AlSkaif, I. Lampropoulos, M. van den Broek, W. van Sark, Gamification-based framework for engagement of residential customers in energy applications, *Energy Research & Social Science* 44 (2018) 187 – 195 (2018). doi:<https://doi.org/10.1016/j.erss.2018.04.043>. URL <http://www.sciencedirect.com/science/article/pii/S2214629618304420>
- [35] J. Kjeldskov, M. B. Skov, J. Paay, R. Pathmanathan, Using mobile phones to support sustainability: a field study of residential electricity consumption, in: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, ACM, 2012, pp. 2347–2356 (2012).

- [36] A. Gustafsson, M. Bång, M. Svahn, Power explorer: a casual game style for encouraging long term behavior change among teenagers, in: Proceedings of the International Conference on Advances in Computer Entertainment Technology, ACM, 2009, pp. 182–189 (2009).
- [37] A. Gustafsson, C. Katzeff, M. Bang, Evaluation of a pervasive game for domestic energy engagement among teenagers, *Computers in Entertainment (CIE)* 7 (4) (2009) 54 (2009).
- [38] D. Wemyss, R. Castri, V. De Luca, F. Cellina, V. Frick, E. Lobsiger-Kägi, P. G. Bianchi, C. Hertach, T. Kuehn, V. Carabias, Keeping up with the joneses: examining community-level collaborative and competitive game mechanics to enhance household electricity-saving behaviour, in: Proceedings of the 4th European Conference on Behaviour and Energy Efficiency Behave 2016, 2016 (2016).
- [39] H. Allcott, T. Rogers, The short-run and long-run effects of behavioral interventions: Experimental evidence from energy conservation, *American Economic Review* 104 (10) (2014) 3003–37 (2014).
- [40] E. Typhina, Designing eco-apps to engage adult learners, in: 2015 International Conference on Interactive Mobile Communication Technologies and Learning (IMCL), IEEE, 2015, pp. 83–87 (2015).
- [41] A. L. Beck, S. Chitalia, V. Rai, Not so gameful: A critical review of gamification in mobile energy applications, *Energy Research & Social Science* 51 (2019) 32–39 (2019).
- [42] J. Hamari, J. Koivisto, H. Sarsa, Does gamification work?—a literature review of empirical studies on gamification, in: 2014 47th Hawaii international conference on system sciences (HICSS), IEEE, 2014, pp. 3025–3034 (2014).
- [43] Y.-k. Chou, Actionable gamification: Beyond points, badges, and leaderboards, Octalysis Group, 2015 (2015).
- [44] L. Morganti, F. Pallavicini, E. Cadel, A. Candelieri, F. Archetti, F. Mantovani, Gaming for earth: Serious games and gamification to engage consumers in pro-environmental behaviours for energy efficiency, *Energy Research & Social Science* 29 (2017) 95 – 102 (2017). doi:<https://doi.org/10.1016/j.erss.2017.05.001>. URL <http://www.sciencedirect.com/science/article/pii/S2214629617301093>
- [45] A. Ortony, G. L. Clore, A. Collins, The cognitive structure of emotions, university press, Cambridge, 1990 (1990).
- [46] C. A. Smith, P. C. Ellsworth, Patterns of cognitive appraisal in emotion, *Journal of Personality and Social Psychology* 48 (4) (1985) 813–838 (1985).
- [47] L. A. Becker, Effect size (es), Retrieved September 9 (2000) 2007 (2000).

- [48] D. Wemyss, F. Cellina, E. Lobsiger-Kägi, V. de Luca, R. Castri, Does it last? long-term impacts of an app-based behavior change intervention on household electricity savings in switzerland, *Energy Research & Social Science* 47 (2019) 16–27 (2019).
- [49] D. A. Activities, P. That, Improve children’s executive functions, *Current directions in psychological science* 21 (5) (2012) 335–341 (2012). doi:10.1177/0963721412453722.
- [50] D. Tellols, M. López-Sánchez, I. Rodriguez, P. Almajano, Sentient embodied conversational agents: Architecture and evaluation., in: *CCIA*, 2018, pp. 312–321 (2018).
- [51] S. Kujala, T. Walsh, P. Nurkka, M. Crisan, Sentence completion for understanding users and evaluating user experience, *Interacting with Computers* 26 (3) (2013) 238–255 (2013).

Post-print