

EVALUATING USER EXPERIENCE OF EARLY PRODUCT CONCEPTS

Virpi Roto^{1,2}, Heli Rantavuo³, Kaisa Väänänen-Vainio-Mattila^{1,2}

¹ Tampere University of Technology, Human-Centered Technology, Tampere, Finland

² Nokia Research Center, P.O.Box 407, 00045 Nokia Group, Finland

³ Nokia Corporation, P.O.Box 100, 00045 Nokia Group, Finland

ABSTRACT

After conducting explorative user studies in the beginning of product development process, there are often more alternative product concepts than it is possible to develop further. The best ideas for production need to be selected by investigating the technical feasibility and profitability of each product concept, but also by evaluating attractiveness and value of the concept proposals for the target user group. This paper discusses means for and challenges in user experience evaluation in the early phases of product development, when only rough product concept descriptions exist. We investigate two lightweight evaluation methods in more detail: expert evaluation and remote online evaluation, and analyse how they could help in identifying the best concepts from the user experience perspective.

Keywords: User experience, Concept evaluation, Expert evaluation, Remote evaluation, Online evaluation

1 INTRODUCTION

Good user experience (UX) is the goal of most product development projects today. User experience is still often seen as result of user interface (UI) design, but many UX researchers argue that good UX comes from the value and meaning of the product concept itself, whereas the UI is providing the means to interact with the product [1],[2]. The UI may provide extra value from entertaining, playful interactions, or it may ruin the user experience by providing difficult routes to product's key functions (i.e. bad usability). However, the main source of sustaining good UX is the value that the user obtains from experiencing the product. The true value of a product is related to the outcome of interaction (the end) and not the interaction itself (means to an end).

According to the above view, the grounding to good UX is laid already when designing the product concept. In order to select the right concept, we need to evaluate the concept ideas. Understanding user needs and evaluating the designs have been the key part of user-centred design (UCD) process, and designing for good UX leans on the same principles. The nature of the work, however, has been shifting from making the user interface easy to use towards promoting more emotional or experiential aspects of the whole product, such as fun, relaxation, trust, social relatedness or self-expression.

The means to evaluate the experiential aspects of the whole concept idea in the early phases of product development are very different from evaluating usability of the first drafts of the user interface. Experiential evaluation focuses on the potential value of the concept idea itself, so evaluation is possible already when we have a rough product concept in place. Even though there is no user interface or interaction design available, we can still evaluate the value of the anticipated interaction outcome. However, value is subjective and context-dependent, so in early evaluations, it would be important to imagine how the concept idea would fit into participant's own context of living. Relating the evaluation to the context of use also helps to improve the concept. The pragmatic aspects, such as usability and functionality of the product are harder to evaluate without existing user interface design, but it is interesting to investigate how we could study the pragmatic aspects in the early phases of product development already.

In this paper, we will investigate the challenges and means for evaluating what we call the "experiential" aspects of products. By experiential, we refer to the fulfilment of the be-goals in contrast to do-goals [3], that is, the non-instrumental, hedonic, or emotional value and meaning of the product. We will start by examining how to establish a systematic way for UX evaluation in chapter 2

and then discuss our insights about early UX evaluation in practice in chapter 3. We will investigate two evaluation methods in more detail: expert evaluation and remote online evaluation.

2 UX EVALUATION

Before we can start evaluating the experiential aspects of products, we have to understand what we actually mean by evaluating user experience. In order to develop a systematic way to evaluate UX, we have taken four steps:

1. Understand what good UX actually means
2. Define metrics for UX
3. Define the methods for evaluating UX (against the metrics)
4. Integrate UX evaluation into the product development process

2.1 Understand what UX is

Although the term user experience has been there since mid-1990's, there is still no widely approved definition for it. Different parties may understand UX in very different ways, ranging from ease of use to immersion, from a fancy user interface to an attitude towards a company. It is thus important to make it clear what we actually mean with UX.

We have investigated user experience by investigating UX literature, by participating scientific workshops discussing the essence of UX, and by surveying how UX professionals see the nature of UX [4]. From these sources, we learned that UX work about is designing for pleasure and value, in contrast to eliminating usability problems. UX is a subjective and holistic concept that is hard to evaluate, since at each moment and for each user, the user experience is different. UX is affected by user's internal state, context of use, and the system used. We also learned that UX focuses on interacting with a product, system, object, or service.

The International Standardization Organization (ISO) is in the process of defining user experience. According to the current version, UX is defined as "a person's perceptions and responses that result from the use or anticipated use of a product, system or service" [5]. The ISO definition is hopefully helping to gain a common understanding about what UX means.

2.2 Define metrics for UX

Since UX is a subjective and holistic concept, it is not easy to define criteria against which it could be evaluated. There are UX evaluations where the participant evaluates the product against the personal criteria [6]. This way of metrics setting is very interesting as it well addresses the subjective nature of UX. It is somewhat harder, however, to compare the results of different subjects, not to mention comparing the results from different product evaluations.

From product creation perspective, each product aims at certain user experience, e.g. fun, trusted, or relaxing. In this case, it is useful to define metrics from the product's perspective rather than from each individual's perspective. The product-centric UX metrics can be universal, that is, applicable for all kinds of products, or specific to a company, brand, or even to one product. On universal level, UX researchers seem to agree that UX has both pragmatic and hedonic (or utilitarian and emotional) aspects [1],[7],[8]. There is also good level of agreement that the pragmatic side includes two key metrics: utility (or functionality) and usability. The experiential (or emotional) side is much more fragmented, as even a broad term such as enjoyment is not always the goal of UX work, e.g. in a security system. We have concluded that each company or brand should define their own goals for the emotional side, as we have done for Nokia [9]. In addition to company-wide UX elements, there are typically product-specific UX targets as well. The more specific the product, the more concrete UX metrics we can define.

The UX targets described above should be taken into account when designing for good UX and they should be transformed as metrics for UX evaluations. For example, in Nokia, one of the elements on the emotional side of UX is Social value, which reflects the company's brand promise of Connecting People. We need to define the metrics for social value by investigating its different aspects. We have identified two major metrics under Social value: feeling connected to others, and identification with the product.

2.3 Define UX evaluation methods

Once we know what the UX metrics are, we can start to define the methodology for collecting the needed data behind the metrics. There are lots of potential criteria for applicable methods [10], so it is clear that one method will not serve all purposes. A set of methods should form a “UX evaluation toolkit” so that UX practitioners can select an applicable method for each case.

Some products and pieces of art are aiming at evoking emotions in the user, as games or movies do. There are methods developed for examining user’s momentary emotions during interacting with a system [11],[12] or for analyzing the emotions after interaction [13]. In product development, finding out the users’ emotions is not enough to improve a product, however, but we need to understand what actually caused the particular emotion. It is quite laborious to run the emotion evaluations and to analyse the reasons behind the emotions.

In UX literature, we have noticed a movement from emotion assessment towards UX of a longer period of time [14][15]. This movement means also a change in the way we evaluate UX: we should not only investigate momentary emotions but also examine how users experience a product as a whole during a longer period. This thinking is well in line with the goals of UX work in companies: make the users pleased with the products and the company so that they will continue using the products of the same company also in the future. The movement towards the longer term UX also suits a broader set of products than the focus in emotion, since not all products are aiming at certain momentary emotions but rather many kinds of positive experiences. We need UX evaluation methods that can tackle the product UX as a whole.

AttrakDiff [16] is an example of a tool suitable for examining long-term UX. It is a well-grounded UX evaluation tool that investigates what adjectives or attributes a user assigns for the examined product. The evaluation tool is simple: the user is presented a questionnaire to evaluate the attractiveness, stimulation, identification, and pragmatic quality of a product with semantic differentials, that is, bipolar adjective pairs such as ugly vs. attractive; inventive vs. conventional; stylish vs. tacky; simple vs. complicated. The questionnaire result shows the hedonic and pragmatic nature of the product, and comparing the results of different products or concept ideas is relatively easy. AttrakDiff questionnaire can be used in many kinds of UX evaluations, and there exist different versions for different types of products¹.

2.4 Integrate UX evaluation into product development

The basic user-centred design principles [17] apply for UX design as well. We should first understand users’ needs and wants in selected contexts, then iteratively design and evaluate the concepts and prototypes during the product development process. The methods are different in different phases of product development: in the very early phases, concept ideas can be evaluated with online surveys or expert evaluations. Before functional prototypes are available, we can run the evaluations in a lab, but once the prototypes are stable and functional enough, a field study is recommended to investigate user experience in the real contexts of use. We can use experts or real users in the early evaluations, as we will describe in the next chapter. Resource-wise it is often good to conduct a lightweight evaluation first, improve the design, then run a lab study with users, and improve the design again. A field study requires the most resources, so it is good to ensure with online surveys, expert evaluations and lab tests that the prototype does not contain any basic or critical problems in the field study.

It is often hard to introduce new activities into the product development process, since processes aim to cut down product development time and adding a new activity often slows down the process. Resource-efficient UX evaluation methods help introducing user studies as new activities into the process in companies that have not followed user-centred design earlier. Also companies that have user studies already built into the product development benefit from finding more lightweight UX evaluation methods, since it will cut down the needed time and resources. Naturally, one needs to analyse the cost of UX evaluation against the gained benefits, and sometimes a heavier evaluation will bring more benefit. Although the development process would slow down a bit in the early phases, it will pay back when the team does not have to make expensive corrections at the later phases. The goal of UX evaluations is to ensure that all products will be valuable and enjoyable for the target users, and this pays back in customer loyalty.

¹ <http://www.attrakdiff.de/en/Home/>

3 EVALUATING EARLY CONCEPT IDEAS

There are many types of evaluation methods available for the early concept ideas. We can use experts or target users as evaluators. We can invite participants to attend a group, pair, or individual session, or ask them to fill in an evaluation questionnaire. There are also many ways to prepare the materials to be evaluated. Sometimes, we can create sketches or mock-ups of the design idea, sometimes a textual or verbal explanation of the idea works better. Contextual information around the concept description would often help evaluators to understand the potential value of the idea, but the risk is that the described context is not relevant for the evaluator her/himself. Comic strips are an interesting combination of sketching, textual explanation, and contextual cues.

In this section, we will investigate two methods in more detail: expert evaluation and a remote online questionnaire. Once the metrics, questions, survey tools and expert/user databases are in place, both of these methods are relatively lightweight to use during a product development process. While these methods have been used extensively for basic usability evaluations, we have not found previous research on using these methods in experiential evaluations. We picked expert evaluation for investigation, because it can be questioned if it is feasible to use experts for evaluating something as subjective as user experience. Online questionnaires are interesting because they enable quick and cost-efficient concept evaluation, also remotely, so that concepts can be evaluated early in different use contexts in different parts of the world. Below, we will discuss the pros and cons of these methods and propose ways to improve their reliability in experiential evaluation.

3.1 Main challenges in early evaluation

A prerequisite for early concept evaluation is to have one or more initial concept descriptions, for example in the form of visual scenarios, draft UI specifications, physical mock-ups – or a combination of these. It is useful to have several alternative concept descriptions prepared for the evaluation, since comparing different ideas is easier for evaluators than providing feedback for just one idea. It is often laborious, however, to prepare such descriptions that are easy to understand for the target users and to ensure that the quality of all descriptions is on the same level. The danger is that an excellent concept idea suffers from unclear or dull presentation and evaluators prefer a less promising concept because of its appealing presentation. The best solution is to have the same professional designer to prepare all concept descriptions. The effort invested in creating clear and attractive concept descriptions pays back not only in the evaluation but also when communicating the idea to the stakeholders inside and outside the company.

In the basic description of the concept idea, not all aspects of interaction are defined, nor can the evaluators “use” the product in real contexts of use. This poses some challenges to the UX evaluation of early concepts. First, the lack of interaction with the intended product leads to incomplete experience of the product. The means to overcome this lack is either to offer the evaluators a demonstrator of the main features and UI elements with some level of interactivity, or to explain how the system would work; the main steps in getting things done. Second, UX is heavily affected by the context of use. The lack of context in early evaluation “strips” many of the contextual elements, especially those relating to physical and social context. As a partial solution, richly described scenarios can set the scene and enable experts to imagine the context-dependent effects to the experience.

3.2 Early expert evaluation

Basic purpose of expert evaluation

One of the central evaluation methods in user-centred design is the expert evaluation, most often conducted as a heuristic evaluation [18]. In this approach, experts – who are experienced either in usability or in the task domain of the evaluated product – will systematically inspect the user interface by using the functionality of the product. The heuristic evaluation is supported by a list of heuristics or rules that describe the recognised usability principles. A well known set of usability heuristics is Jacob Nielsen’s 10 usability heuristics [19], which include heuristics for *visibility of system status*, *user control and freedom*, *error prevention*, for example. Because expert evaluation can be done in a relatively short time frame, and no end-users are involved, expert evaluation is often referred to as a discount usability engineering method.

Expert evaluations are most needed for concepts that are targeting distant future and rely on having other technological advancements in place. The risk of evaluating future concepts with users is that

they might not see the relevance of a concept until they have experienced a clear need for it, and thus abandon the futuristic concepts. For example, before the social networking sites became popular in the Internet, ordinary users might not have seen it very tempting to be able to access this kind of information on a mobile phone. Experts who saw this trend coming could see the value even before this behaviour became commonplace. Therefore, experts who have insight about how the world is likely to change before the evaluated concept product will come out are sometimes the best ones to evaluate futuristic concepts.

When there are tens of different concept ideas to choose from, it requires effort to produce concept descriptions that would be easy to understand by ordinary users. Some concept ideas are so complicated that they would themselves require an iterative design-evaluation cycle to get the description clear enough for laymen. Since experts are often able to understand the idea even without an elaborate concept description, expert evaluation can help when it is too laborious to generate high quality concept descriptions.

Expert evaluation of user experience

The background of heuristic evaluation was on PC-based graphical user interfaces, but usability heuristics have been extended in many studies [20] to better suit specific domains, for example, Web sites [21], playability or gaming experience [22] and social interaction [23]. Some of these approaches are expanding to a broader scope beyond usability and try to establish ways to do early evaluations of experiential aspects of using a product.

UX is challenging to evaluate by experts because experiences are personal by definition. Thus, as a starting point, experts may be able to evaluate the product from their own perspective only, and not from a broader user populations' perspective. Furthermore, the commonly known usability heuristics do not cover the broader experiential aspects of product usage. In their study on user experience of Web2.0 type of services, [24] developed a set of UX evaluation heuristics that cover both pragmatic and hedonic aspects of Web service UX. These heuristics were specific to Web 2.0 services and covered issues such as fluent cross-platform service usage, social interaction and navigation, dynamic content and UI update, and context-awareness. For each of the heuristics, pragmatic and hedonic aspects were described. An example is presented in Table 1.

Service UX evaluation heuristic (example): Dynamic service features	
<i>Description:</i> Services are “live” entities because of their dynamic nature. Services change constantly on various levels: the service content (provided by service providers and/or generated by users), their user interface, or even their functionality may change over time. For the user, the main challenge is to maintain an overview of “what is there”, “what is new” and “what is interesting to me”. The opportunity of the temporal nature of services is that users will be excited about the constant novelty and timeliness of service contents and the UI.	
Pragmatic aspects of the sample heuristic	Hedonic aspects of the sample heuristic
<ul style="list-style-type: none"> • When users enter the service, it is possible to gain an overview of the recent changes in the service • While using the service, users can easily find the new content that is interesting to them 	<ul style="list-style-type: none"> • The service feels like a lively place where it is enjoyable to spend time • The service satisfies users' curiosity/seeking of knowledge by frequently offering interesting content • The changes in the UI or functionality are seen as improvements to the service

Table 1: Pragmatic and hedonic aspects of a service UX heuristic [24]

To test the applicability of heuristics in expert evaluation of UX, we tested three Web 2.0 services: Facebook, Nokia Sports Tracker, and Trip Advisor. In the actual expert evaluation process, each of the three services was evaluated by three UX experts for about 10 hours each. The evaluators had 1-10 years of experience in usability and user experience design and evaluation. The UX experts used the services and listed all their findings related to the heuristics. They were also asked to identify their findings under “positive” and “negative” findings, and under “pragmatic” or “hedonic” categories. The results of the evaluation indicated that experts were able to identify experiential aspects of usage, both positive (36% of the findings) and negative (64%), from both pragmatic (74% of the findings) and hedonic (26%) viewpoints. The evaluation results gave a broad picture of UX of each of the evaluated

services. In addition, new information was compiled about the detailed aspects of service UX, for example, on the importance of expressing one's achievements to other users, and on informing users about the changes in the service. The evaluators expressed some concerns about the adequacy of the length of the evaluation period – for example the social interaction would require a relatively long time to gain in-depth experiences. Some of the evaluators also hoped to have been offered scenarios which would have helped them to focus on the central issues of use. [24]

We argue that expert evaluations can give valuable information of UX aspects of products even on the early conceptual level. Naturally, in addition to the experts, a broader set of potential users should be involved in evaluation in selected phases of the design process. However, during the various stages of concept design iterations UX expert evaluations can be used as a valuable discount evaluation method to keep the design on the right track.

3.3 Remote online evaluation with users

In early product concepting, user experience evaluation is typically carried out face to face with a group of potential users of the concept. Arranging and carrying out face to face evaluation sessions with participants requires particular resources that range from user study methodology expertise to time, which are not always available to all teams that would benefit from end user involvement. In addition, face to face evaluation sessions are typically only possible to arrange within the same geographic location where the design team is situated, although the product that was being designed was targeted at consumers at locations globally.

Therefore, to complement face to face evaluation, Nokia has developed an online platform called Nokia Pilots for involving users in product development. The platform has been developed in close collaboration with Nokia Design and Nokia Research Center, utilizing the latest knowledge of user-centred design methodologies and user experience research internationally. Nokia Pilots online platform engages end users into product development through a methodology that applies the Nokia user experience elements [9] and internationally established user-centred design methods. Regarding early concept evaluation, the platform utilizes the fact that in product development, early concept ideas are often formulated as high level scenarios that examine the intended product by describing its context of use, using media files that combine text, image and audio. The high-level scenarios represented as media files are well suited for online evaluation by users when compared to some other concept types: evaluating physical prototypes, for example, requires observation of the users' tactile experience and interviewing them face to face. Early high-level scenarios invite emotional, experiential reactions on the value and meaning of the product concept, and it is possible to explicate these reactions in writing or express them through questionnaire answers. Secondly, it is possible to deliver the media files quickly on the Web to various geographical and cultural contexts. Remote evaluation has its limitations as a method for experiential evaluation, which is discussed below, and as mentioned in section 2.3, is best used as part of a "UX evaluation toolkit." However, for a global company, the method can gain key importance as it reaches users from all market areas cost-efficiently and enables quick concept iterations based on their feedback.

One of the early pilots to test the platform was a design project focused on accessories to mobile devices. The project was at a stage where the team was in need of user feedback for their early concepts, in order to know how to modify them towards the experience that the users desire. The team had a long experience in carrying out face to face user evaluation workshops to gather experiential user feedback. Their main interest for using online evaluation was to be able to engage users from varied locations around the world. It was also of interest to the project to see how much time and human resources they would be able to save, but still collect high quality results.

The participants for the Nokia Pilots studies are recruited through the public Nokia Pilots registration page on the Web. Product development projects that wish to invite Nokia Pilots users to participate in their program search the Nokia Pilots user database to form a suitable test group. They may form either random representative samples or samples that in qualitative research methodology are called non-random purposive or theoretical samples, commonly used in exploratory research [25]. Purposive samples are typically smaller than random ones, and in data produced, the emphasis is on qualitative data.

In purposive sampling, respondents are selected because they belong to a special population that is especially informative considering the topic on which information is needed. For example, in a study that aims to know how camera phone photos are used, it is not very useful to interview users who have

never used camera phones. In the case of the accessories study, the most significant criteria were that the respondents had some experience in the field that the accessory was related to and that they came from the intended market areas. It is also relevant to note that the respondent could withdraw from participating in the project at any moment.

The respondents on different continents were forwarded five storyboards and a questionnaire online. The storyboards represented different use cases of the product concept. The questionnaires related to the storyboards and to the respondents' habits and practices regarding the area of the designed accessory. The storyboards were similar to comic strips. Each consisted of a few simple drawings with text to explain what took place and what was relevant in the drawing. Figure 1, for example, was the first in a sequence of four pictures. Each storyboard represented a particular use case related to an accessory. The drawings illustrated the physical and social contexts where the accessory could be used. Particular functions and user interface views were highlighted, at the level of basic functionality, to be evaluated within this context specifically. The questionnaire included, first, a section for general questions on the respondent's usual practices related to the planned context of use. Next, questions were presented related to each use case. The questions invited subjective evaluations of the use cases presented in the storyboards, and the questions also presented alternative use cases which the respondent could show preference for. The answer options were multiple or single choice and free text. In multiple or single choice answer options, the respondent ticked one or several answers from a list of alternatives written by the project staff. No numerical scales were used. With the storyboards and the questions related to them, the project sought to validate their initial assumptions of the contexts in which the new products would be used, and the desired functionalities for the product. The target was to collect feedback about the social, emotional, and utility aspects of the product concepts, rather than about the appearance and functional aspects of them. The users had five days to review the storyboards and to fill in and submit the questionnaires.



John is sitting in his car on his way home from work

Figure 1. Part of a storyboard presenting a concept idea

Pros and cons of remote online evaluation

The project learned that a good questionnaire design is a prerequisite for results that meet the concepting project's needs. Thorough questionnaire design takes time and expertise. However, it is necessary in order to direct the feedback from users towards the intended items and questions, and at the same time to encourage the respondents to express their own ideas from outside of the questionnaire structure. Results from a well-devised questionnaire provide actionable answers to the open questions that the project has as well as new, unanticipated information, are valid, and easy to collate and analyse. Especially, the project became aware of the risk that the respondents may not fully understand the use cases or the questions related to them, as there would be no-one to turn to for questions as in face to face situations. Compared with face to face evaluation of concepts with end users, online questionnaires present a higher risk of low engagement with the concepting materials and the questionnaires. To reduce the risk of low involvement that may result in low quality data (random or incomplete answers, high percentage of non-response), questionnaires are best to be formulated as concise and consistent, and a manageable number of concepts should be presented for the user. At the same time, it is useful to note that in remote evaluation, the respondent can distribute the time that she or he spends on evaluating each concept more freely than in face to face workshops. With the materials online, the respondent can choose when to evaluate the concepts and how much time to take for the evaluation.

The global reach of an online user database poses specific requirements to both the concepting materials and questionnaire design. The designer of the evaluation materials must be careful not to distract the user in a different cultural sphere with anything that she or he may not understand or that may gain ill-proportioned significance compared to its intended role – for example, colour codes or form factors.

At the same time, the global reach is where the value of online concept evaluation lies: to bring out cultural issues at a very early stage in product development. The project in our case wanted to examine three themes especially: national rules and regulations, values that users had related to style and fashion, and values that they had related to privacy. They picked respondents from countries where, according to the understanding within the project, there were few rules and regulations concerning the use context, and where the context was highly regulated. Some of the use cases that the project exposed to respondents were written with the awareness and expectation that they would spark stronger reactions than others based on the different rules and values across cultures. The project found this strategy successful: based on the user feedback that included evaluation and suggestions, the project was able to focus the next creative session to deal with the cultural particularities. As this example shows, for these issues to surface from the data there must be competence in interpreting the replies so that the role of cultural factors are understood instead of leaving them unnoticed or labelling them as individual variance.

With online questionnaires, although they can also be used to obtain qualitative data, it is difficult to reach the depth of qualitative understanding that is achieved in face to face encounters. This was recognised also by our case project. However, with cultural factors in mind especially, online feedback has the benefit of being less intimate than feedback given face to face. The respondent can be assumed to be able to be more candid and honest than a workshop participant who gives her feedback personally to the moderator in a group situation. Especially in cases where the evaluation concerns a social aspect, for example, whether the product would cause unwanted social consequences or emotions, the online respondent may experience it as easier to disclose these concerns online than in face to face interaction. Online, the respondent also has fewer cues to interpret what kind of reaction is expected from her than in face to face situations. Considering these factors is crucial in planning experiential evaluation by users especially as the emotional aspects of UX are, as mentioned above, are becoming more and more important in product design.

Another challenge in global online evaluation is how to best reward the participants. In face to face workshops, the social interaction that takes place in the encounter may form a part of a rewarding experience. Online, the same can be attempted by making use of social media techniques, such as discussion forums or instant messaging, through which the respondents can interact with each other, the projects, and the people who sustain feedback platform. From the methodological point of view, social media techniques also allow to expand the collection of qualitative data in remote evaluation outside the questionnaire format. However, a global platform brings in users with very different backgrounds and interests, and it must be acknowledged that not all users share an interest in participating in web discussions or networking with other users. Rewards should also be compliant with each of the cultural contexts that the users are located in.

The global scale of the user database and the frequent need for small, purposive samples pose requirements for rich profile data and efficient search functions to find the relevant participants for each project. Scarce profile data would cause pressure for verifying the samples with further questions, which wastes time and effort of both the project and the user, and it is not pleasant for the user to become rejected in the sample verification stage. Also, in the absence of an efficient filtering function in creating the sample erodes the efficiency benefits of the online evaluation.

From the company perspective, one of the challenges of early product evaluation online is how to safeguard new product ideas and strategic directions from competitors, while being receptive to end users' input to these ideas and strategies, and thereby, to future products. Understanding of the type and magnitude of this challenge will cumulate as more cases, such as the accessory design project in this case, are carried out and more experience is gained from user engagement online. However, at present, there are already several measures for safeguarding product ideas while exposing them for evaluation. Paying close attention to how the evaluation materials are formulated; confirming the identity of the respondents; and using non-disclosure agreements when necessary are among the basic ones. Overall, the benefits of involving users in product development online still currently outweigh the potential challenges involved. The online environment presents a possibility to bring users and product developers into direct contact, with increased possibilities to detect new product ideas directly from the user feedback.

Different design teams work with varying targets, resources, and varying methods of involving users and experiential evaluation in the design process. An online platform that provides a standard set of methods, which do not require much resources to apply, can introduce new methods of engaging users

to concepting teams. Such a platform can also introduce user evaluation as a way of working to teams who have faced difficulties with finding the time and people for this work before, and thus strengthen the user-centred approach within the company. At the same time, it is important that the online platform is flexible to the different needs of the concepting teams. In designing the platform itself, this means a fine balance between the attempt to set a standard for online user evaluation methods through building the online service on best practices, and allowing for variety within those practices to further evolve.

Above, some challenges and requirements have been discussed that concern the users, the project who invites them to participate, and the platform on which the respondents and projects work on. From the company point of view, other issues emerge that are relevant to consider. At the company level, the automated collection of data that is possible with online evaluation holds many promises. It helps to accumulate historical data on how different concepts have been evaluated by users, and it is possible to make results from previous project easily accessible by others later on. This way, it is possible to accumulate knowledge in certain product fields on user reactions to concepts in this field, and also to gradually learn about the best practices of online user engagement in early product evaluations.

4 CONCLUSIONS AND FUTURE WORK

In this paper, we have investigated the early phases of designing for pleasure with the focus on evaluating the experiential aspects of concept ideas. We described the steps we have taken in order to establish a systematic way to evaluate UX: Gaining common understanding of UX, developing metrics and methods, and fitting the evaluations into the product development process. These steps were needed since the whole UX research is still a new field and there are no established best practices for evaluating the experiential aspects of UX. This paper provides both the theoretical background and an empirical case for showing how UX evaluation can be arranged in the early phase of product development.

In industry setting, UX evaluation methods need to be fast and lightweight, otherwise it is difficult to get resources to evaluate UX often enough and in several locations around the world. In this paper, we investigated two relatively lightweight evaluation methods, suitable also for concept evaluation: Expert evaluation and online questionnaires. Both methods need some effort before they can be taken into a systematic, routine use, but once the tools are in place, studies can be done even within a day.

In the early evaluation, quite some effort needs to be invested in preparing equally attractive and clear enough concept descriptions for a fair evaluation. Expert evaluation can be used for gathering quick feedback on design ideas that are not yet articulated clearly. Experts can also help in evaluating concepts that are targeting distant future and are dependent on other technological advances that ordinary people are not aware of. The main challenge for the experts is to foresee how ordinary users will experience the concept and to be able to evaluate something as subjective as UX from expert point of view.

The other UX evaluation method investigated in this paper was an online questionnaire. The main benefit of online UX evaluation is to reach participants from various locations with small effort. An online questionnaire also provides data that is typically quicker to analyse than interviews. We discussed the challenges that online UX evaluation poses, for example, collecting rich user profiles to the user database, improving low engagement with the evaluation, and rewarding respondents.

The field of UX research is just emerging, and there is a lot of interesting work to be done in developing practical UX evaluation system with metrics, methods, and tools applicable for various needs and situations. For early concept evaluation, we are interested in developing best practices for creating evaluation materials, since good concept descriptions are elementary for reliable UX evaluation in the early phases of product development. We need to develop UX metrics and UX targets applicable all the way from concept ideas to ready products. We are also eager to see new kinds of experiential evaluation methods, since there is a lot of room for innovation in the ways of evaluating UX.

REFERENCES

- [1] Hassenzahl, M. The Thing and I: understanding the relationship between user and product. In Blythe, M.A., Overbeeke, K., Monk, A.F., Wright, P.C. (Eds.), *Funology: From Usability To Enjoyment*, 2003, pp. 31-42 (Kluwer Academic Publishers, Norwell, MA).

- [2] Cockton, G. Putting value into E-valuation. In Law E., Hvannberg, E., Cockton, G. (Eds.), *Maturing Usability*, 2008 (Springer-Verlag London Limited).
- [3] Carver, C. S., & Scheier, M. F. *On the self-regulation of behavior*, 1998 (New York: Cambridge University Press).
- [4] Law, E., Roto, V., Hassenzahl, M., Vermeeren, A., Kort, J. Understanding, Scoping and Defining User eXperience: A Survey Approach. In *Proceedings of CHI'09*, Boston, MA, USA, April 2009 (ACM).
- [5] ISO DIS 9241-210:2008. *Ergonomics of human system interaction - Part 210: Human-centred design for interactive systems* (formerly known as 13407). International Organization for Standardization (ISO). Switzerland.
- [6] Karapanos, E., Martens, J.-B. and Hassenzahl, M., Accounting for diversity in subjective judgments, In *Proceedings of CHI'08*, Florence, Italy, April 2008 (ACM).
- [7] Jordan, P. *Designing Pleasurable Products: An introduction to the new human factors*, 2002.
- [8] Norman, D. *Emotional Design: Why We Love (or Hate) Everyday Things*. 2004, Basic Books, USA.
- [9] Roto, V., Rautava, M.: User Experience Elements and Brand Promise. *International Engagability & Design Conference (Idec4)*, in conjunction with NordiCHI'08 conference. October 19, 2008, Lund, Sweden (2008)
- [10] Väänänen-Vainio-Mattila, K., Roto, V., Hassenzahl, M.: Towards Practical User Experience Evaluation Methods. 5th COST294-MAUSE Open Workshop on Valid Useful User Experience Measurement (VUUM). Reykjavik, Island (2008)
- [11] Mandryk, R.L., Inkpen, K., and Calvert, T.W. Using psychophysiological techniques to measure user experience with entertainment technologies. *Behaviour and Information Technology* (Special Issue on User Experience), 25,2, (2006), pp.141-158. (Taylor & Francis).
- [12] Isbister, K., Höök, K., Sharp, M., and Laaksolahti, J. The sensual evaluation instrument: developing an affective evaluation tool. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Montréal, Québec, Canada, April 22 - 27, 2006) CHI '06 (ACM)
- [13] Desmet, P. M. A. Measuring emotions. In M. Blythe, C. Overbeeke, A. F. Monk, & P. C. Wright (Eds.), *Funology: From Usability to Enjoyment*. 2003, Dordrecht: Kluwer.
- [14] Norman, D. A. THE WAY I SEE IT - Memory is more important than actuality. *Interactions* 16, 2 (Mar. 2009), pp. 24-26.
- [15] Karapanos, E., Zimmerman, J., Forlizzi, J., and Martens, J. User experience over time: an initial framework. In *Proceedings of the 27th international Conference on Human Factors in Computing Systems* (Boston, MA, USA, April 04 - 09, 2009). CHI '09 (ACM).
- [16] Hassenzahl, M., Burmester, M., & Koller, F. 2003a. AttrakDiff: Ein Fragebogen zur Messung wahrgenommener hedonischer und pragmatischer Qualität. In J.Ziegler & G. Szwillus (Eds.), *Mensch & Computer 2003. Interaktion in Bewegung* (pp. 187-196). Stuttgart, Leipzig: B.G. Teubner.
- [17] ISO 13407:1999. *Human-centred design processes for interactive systems*. International Organization for Standardization (ISO). Switzerland.
- [18] Nielsen, J. and Mack, R.L. (Eds). *Usability Inspection Methods*, 1994 (John Wiley & Sons Inc.).
- [19] Nielsen, J. Heuristic evaluation. In Nielsen, J., and Mack, R.L. (Eds.), *Usability Inspection Methods*, 1994 (John Wiley & Sons, New York, NY).
- [20] Ling, C., Salvendy, G. Extension of heuristic evaluation method: a review and reappraisal. *Ergonomia IJE&HF*, Vol. 27, No. 3, (2005) pp. 179-197.
- [21] Nielsen, J. *Designing Web Usability*, 1999 (Peachpitt Press).
- [22] Korhonen, H. and Koivisto, E. Playability heuristics for mobile games. In *Proceedings of MobileHCI'06*, 2006, pp. 9-16 (ACM).
- [23] Preece, J. *Online Communities: Designing Usability, Supporting Sociability*, 2000 (John Wiley & Sons, UK).
- [24] Väänänen-Vainio-Mattila, K. and Wäljas, M. Development of Evaluation Heuristics for Web Service User Experience, in *Adjunct Proceedings of CHI'09*, Boston, MA, USA, April 2009 (ACM).
- [25] Neuman, D. *Social research methods: Qualitative and quantitative approaches*, 2006 (Pearson)