Interacting with Computers xxx (2011) xxx–xxx

Contents lists available at ScienceDirect

Interacting with Computers

journal homepage: www.elsevier.com/locate/intcom

UX Curve: A method for evaluating long-term user experience

Sari Kujala [a,*], Virpi Roto [b], Kaisa Väänänen-Vainio-Mattila [a], Evangelos Karapanos [c], Arto Sinnelä [a]

[a] Tampere University of Technology, P.O. Box 589, FI-33101 Tampere, Finland
[b] Nokia Research Center, P.O. Box 407, FI-00045 Nokia Group, Finland
[c] Eindhoven University of Technology, Department of Industrial Design, P.O. Box 513, 5600 MB, The Netherlands

ARTICLE INFO

Article history:
Received 10 September 2010
Received in revised form 30 June 2011
Accepted 30 June 2011
Available online xxxx

Keywords:
User experience evaluation
Long-term user experience
Mobile phone
User satisfaction
Recommendation

ABSTRACT

The goal of user experience design in industry is to improve customer satisfaction and loyalty through the utility, ease of use, and pleasure provided in the interaction with a product. So far, user experience studies have mostly focused on short-term evaluations and consequently on aspects relating to the initial adoption of new product designs. Nevertheless, the relationship between the user and the product evolves over long periods of time and the relevance of prolonged use for market success has been recently highlighted. In this paper, we argue for the cost-effective elicitation of longitudinal user experience data. We propose a method called the "UX Curve" which aims at assisting users in retrospectively reporting how and why their experience with a product has changed over time. The usefulness of the UX Curve method was assessed in a qualitative study with 20 mobile phone users. In particular, we investigated how users' specific memories of their experiences with their mobile phones guide their behavior and their willingness to recommend the product to others. The results suggest that the UX Curve method enables users and researchers to determine the quality of long-term user experience and the influences that improve user experience over time or cause it to deteriorate. The method provided rich qualitative data and we found that an improving trend of perceived *attractiveness* of mobile phones was related to user satisfaction and willingness to recommend their phone to friends. This highlights that sustaining perceived attractiveness can be a differentiating factor in the user acceptance of personal interactive products such as mobile phones. The study suggests that the proposed method can be used as a straightforward tool for understanding the reasons why user experience improves or worsens in long-term product use and how these reasons relate to customer loyalty.

1. Introduction

User experience is a multidimensional concept and a commonly accepted definition is still lacking. As Hassenzahl and Tractinsky (2006) argue, the concept of user experience attempts to go beyond the task-oriented approach of traditional HCI by bringing out aspects such as beauty, fun, pleasure, and personal growth that satisfy general human needs but have little instrumental value. Therefore, when compared to basic usability, *enjoyability* plays an essential role in user experience. The extent to which an interactive product is enjoyable to use is referred to as the product's hedonic quality (Hassenzahl et al., 2000; Hassenzahl, 2001). The shift of emphasis from usability to experiential factors has forced researchers to consider what user experience actually is and how to evaluate it (Vermeeren et al., 2010).

A number of user experience models have been proposed during the past decade (Desmet and Hekkert, 2007; Forlizzi and Battarbee, 2004; Hartmann et al., 2008; Hassenzahl, 2004; Hassenzahl and Tractinsky, 2006; Jordan, 2000). Mahlke and Thüring's (2007) holistic model is one of the most recent and it is building on the previous models and research findings. Mahlke and Thüring (2007) distinguish three user experience components which together determine the user's overall appraisal of a system and influence their future decisions and behavior: perception of instrumental qualities, emotional reactions and perception of non-instrumental qualities. Instrumental qualities are related to the usability and usefulness of a system, while non-instrumental qualities are related to its appeal and attractiveness. Mahlke and Thüring (2007) also found empirical evidence that both of these two aspects of quality influence emotional reactions and the appraisal of interactive systems.

It has been found that 'perceived usefulness' loses its dominant predictive power when applied to hedonic systems such as a movie website (Van der Heijden, 2004), and this is where perceived enjoyment becomes useful as a factor for explaining user acceptance. Schrepp et al. (2006) show that hedonic quality has also an important impact on the attractiveness of business management software that is used for work purposes. Companies see good

* Corresponding author. Tel.: +358 50 3862768; fax: +358 3 31155532.
E-mail address: sari.kujala@tut.fi (S. Kujala).

0953-5438/$ - see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.intcom.2011.06.005

user experience as being vital for continuous commercial success, especially since positive long-term user experience is believed to improve customer loyalty. For example, Garrett (2006) and Jordan (1998) suggest that if people have delightful experiences with a product, they are more willing to buy the next product from the same company. Indeed, the results of Chitturi et al. (2008) show that delighting customers with hedonic quality improves customer loyalty in car owners, as measured by word of mouth and repurchase intentions, more than utilitarian quality does.

If enjoyment of hedonic quality plays an essential role not only in user experience but also in creating customer loyalty, the question then arises of what the appropriate timeframe for evaluating user experience is. Traditional usability evaluation methods focus on 'first-time' experiences with products and learnability (Courage et al., 2009; Mendosa and Novick, 2005). Mendosa and Novick (2005) followed middle-school teachers creating Web sites for 8 weeks and found in their longitudinal study that the type of errors and the causes of the users' frustrations changed radically over time. Thus, the errors found in usability testing may have little consequence over longer periods of time. Similarly, learning has an effect on the hedonic aspects of use, but not much has been studied as to how user experience evolves over time (Karapanos et al., 2009b; von Wilamowitz-Moellendorff et al., 2006).

Karapanos et al. (2008) found that during the first experiences of using a novel pointing device pragmatic aspects were the most prominent determinant of goodness, but after 4 weeks of use identification (i.e. what the product expresses about its owner) became a more prominent aspect of the product. On the other hand, stimulation lost its power to make a novel pointing device beautiful over the first 4 weeks of using and owning the product. Similarly, Karapanos et al. (2009b) followed six participants for 1 month after the purchase of an Apple iPhone and found that, while the importance of its novelty and social meanings quickly faded away, over time different sources of hedonic quality emerged such as the participation of the product in daily rituals and cherished activities that become habituated. von Wilamowitz-Moellendorff et al. (2006) found that the perceived stimulation and other hedonic aspects of user experience of mobile phones seem to fade away during the first 20 months of use. In this regard too, Fenko et al. (2010) asked users to describe their sensory experiences with different consumer products and rate the importance of different sensory modalities during the first year of usage; they found that at the moment of purchase, vision is the most important modality, but after 1 month touch and audition became more important.

Although, the time dimension of user experience has been repeatedly noted as being particularly important (Courage et al., 2009; Hassenzahl and Tractinsky, 2006; Karapanos et al., 2009b), most of the current UX evaluation methods focus on single behavioral episodes and momentary evaluations and only 36% of methods focus on long-term period of experience (Vermeeren et al., 2010). There are also few published research studies (Karapanos et al., 2009b; von Wilamowitz-Moellendorff et al., 2006). Even, in the existing longitudinal studies of user experience, the investigated time periods are relatively short, ranging from short test tasks (Hassenzahl et al., 2000; Mahlke, 2006; Tractinsky and Zmiri, 2006) to 5 weeks of usage (Courage et al., 2009). Some studies, such as von Wilamowitz-Moellendorff et al. (2006) and Karapanos et al. (2009a,b, 2010), focus on the evolving nature of user experience over months or even the complete lifecycle of a product; such studies are very rare because of the expense, loss of time, and participant fatigue involved (Karapanos et al., 2009a).

In this study, we present the first results of using a curve drawing method – called the UX Curve – that can be used to investigate various dimensions of UX over time. We evaluated the usefulness of the UX Curve method by investigating whether it supports mobile phone users in producing qualitative data about their experiences. In addition, we analyzed the relationship of those experiences with user satisfaction. A questionnaire was also used to measure overall satisfaction, customer loyalty, willingness to recommend to others and actual recommending.

2. Evaluating user experience over time

Because evaluating a momentary user experience is in most cases not very reliable for predicting user experience in real life or for assessing the success of a product, we need information about *long-term user experience* – how the user's experience and relationship with a product evolves over time from the early learning and enthusiasm to becoming a part of daily life. It is this long-term user experience that makes people continue to use a product and to recommend it to others – not each detail of their individual experiences.

Whilst measuring first impressions and momentary experiences is important for getting feedback in the early development stages (Vermeeren et al., 2010), it is known that the relevance and salience of different user experience aspects changes over time (Fenko et al., 2010; Karapanos et al., 2009b). Creating customer loyalty thus demands more sustained elicitation of feedback over longer periods of time. Such feedback may inform the design of subsequent generations of the same or similar products of a brand. In addition, since user experience is highly dependent on the user's internal state of mind and context of use (Hassenzahl and Tractinsky, 2006), field studies provide a much more realistic context within which to obtain reliable user experience data (Vermeeren et al., 2010). However, field studies are very time-consuming and cumbersome to carry out for longer time periods.

The Experience Sampling Method (ESM) was developed to collect information on people's reported feelings in real time in natural settings during selected moments of the day (Kahneman et al., 2004). ESM minimizes the bias caused by retrospection, but because it burdens participants by asking them to interrupt their current activity, conducting ESM for longer periods of time is very difficult. This motivated Daniel Kahneman and his colleagues to develop the Day Reconstruction Method (DRM) Kahneman et al. (2004). DRM imposes a chronological process in the reconstruction of daily experiences which has been found to increase respondents' ability to reconstruct behaviors and emotions experienced in daily activities. Kahneman et al. (2004) reported results very close to those elicited through the Experience Sampling Method. The benefit of DRM over ESM is that takes less time, does not interrupt respondents' daily activities and imposes a smaller burden upon them (Karapanos et al., 2009b).

Karapanos et al. (2009b) used the Day Reconstruction Method (DRM) to evaluate users' experiences with iPhones for 5 weeks. Their findings were that many different kinds of experiences may take place during the same day, but their distribution changes over time, starting from an orienting learning phase to a final emotional attachment phase. However, as they reported, these kinds of longitudinal study are very rare because of their expense and laborious nature as participants have to sustain motivation in reporting for the full length of the study (Karapanos et al., 2009a).

Furthermore, since the user's overall evaluation of user experience is not a simple sum of the individual experiences (Tractinsky and Zmiri, 2006), we can question whether it is necessary to measure all the individual experiences over time. Users form their overall evaluation of the product on the basis of *memories* of past experiences and psychological studies indicate that people cannot remember all the details of their experiences (Norman, 2009). In addition to losing details, memory introduces systematic biases into evaluations (Hsee and Hastie, 2006). Fredrickson's (2000) review of the psychological literature showed that people base their

Please cite this article in press as: Kujala, S., et al. UX Curve: A method for evaluating long-term user experience. Interact. Comput. (2011), doi:10.1016/j.intcom.2011.06.005

overall evaluations on the peak and final intensity of the experience. In addition, the *chronological order* of the component events and the trend of the experience influence the overall evaluations (Fredrickson, 2000). In particular, people like happy endings and prefer experiences that improve rather than ones that get worse (Ross and Simonson, 1991).

Norman (2009) and Karapanos et al. (2010) argue that *these memories of experiences will be reported to others and guide the future behavior of the individual.* Thus, the reconstructed memories are very relevant, in spite of the possible bias in recalling them. For example, Oishi and Sullivan (2006) show that retrospective evaluations predict later human behavior better than daily ratings do. In addition, longitudinal real-time ratings are often expensive and impractical, especially in product development contexts. To this end, alternative and easier to apply long-term UX evaluation methods founded on retrospection are needed.

von Wilamowitz-Moellendorff et al. (2006) developed a retrospective interview technique named CORPUS to reconstruct changes in UX over a period of more than 1 year. Users were also asked to rank several UX quality dimensions with a 10-point scale on a timeline. The technique was a simple way to identify the dynamics of UX dimensions and the researchers were able to identify two phenomena, i.e. familiarity and comparison with other phones, which explained the changes over time.

In addition, Karapanos et al. (2009a, 2010) developed iScale, a survey tool that aims to assist users to retrospectively recalling their longitudinal user experience with a product by "sketching" a curve over time. In iScale users are asked to sketch how their opinion had changed from the moment of purchase up until the present. One study by Karapanos et al. (2009a) compared the effectiveness of two versions of iScale against free recall without any form of sketching. Compared to its analog version, free-hand sketching, iScale was found to have a number of limitations but also some benefits. Free-hand sketching was found to be more expressive than iScale due to the increased degrees of freedom in sketching as well as giving users the ability to easily annotate sketches. On the other hand, iScale provides the ability to modify the sketch at a later point and it is possible to use it as an online survey tool.

Methods like CORPUS and iScale provide two kinds of data: a recalled pattern of change of user experience and self-reports of personal experiences that induced the changes in their perceptions. These methods were designed to minimize the retrospection bias and the results of Karapanos et al. (2009a) show that the retest reliability of the recalled information is rather good. Neither von Wilamowitz-Moellendorff et al. (2006) nor Karapanos et al. (2009a) included the interpersonal analysis of the sketched graphs however, so it remains unclear how the chronological order of experiences affects the overall evaluation and customer loyalty.

With these influences in mind, the UX Curve was developed as an easy to apply method for supporting users in recalling important details of the product qualities that affect user experience. The UX Curve and iScale (Karapanos et al., 2009a, 2010) were developed independently from each other. While iScale was designed mostly as online survey technique, UX Curve is intended to be used in a face to face setting where the researcher is better able to inquire into the participants' reasoning and thoughts. In addition, the dimensions of user experience investigated are different: the iScale focuses on usefulness, ease-of-use and innovativeness, whereas the UX Curve investigates attractiveness, ease of use, utility, and degree of usage. Furthermore, while Karapanos et al. (2009a) aim at assessing the validity and reliability of recalled information, we attempt to take a step forward by studying the meaningfulness of the recalled information in relation to user satisfaction and customer loyalty. Neither method is useful until it is known what the resulting data means and what separates happy and unhappy users.

3. Developing and piloting the UX Curve method

Our aim was to develop an easy to apply method for evaluating long-term UX. We were particularly interested in identifying the chronological order of experiences as this aspect has been shown in psychological studies to not only affect users' overall evaluations (Fredrickson, 2000), but also to predict later behavior (Oishi and Sullivan, 2006). We developed a template (see Fig. 1) so that users could draw themselves a curve describing how their experiences had evolved over time. The instruction given with the template was: 'Please recall the moment when you began to use the product. Draw a curve describing how your relationship towards the product has changed from the first time you used it until today.' Users were then asked to mark the reasons at their approximate locations on the curve.

The template includes an empty two-dimensional graph area and lines for writing on and briefly describing the reasons for the changes in the curve. The horizontal axis represents the time dimension from the beginning of use to the current moment and the vertical axis represents the intensity of the users' experience. In the middle of the graph area there was a horizontal zero line dividing the area into a positive upper part and a negative lower part. The vertical axis was labeled accordingly with + and − signs.

As user experience is a complex concept and not all of its aspects have been fully established, we decided to let users freely express their experiences with the first general UX Curve. Thus users were able to report the kinds of experiences that they found meaningful. In order to complement the general view, we next asked users to consider their experiences from four other viewpoints, three of which were selected to correspond to the user experience

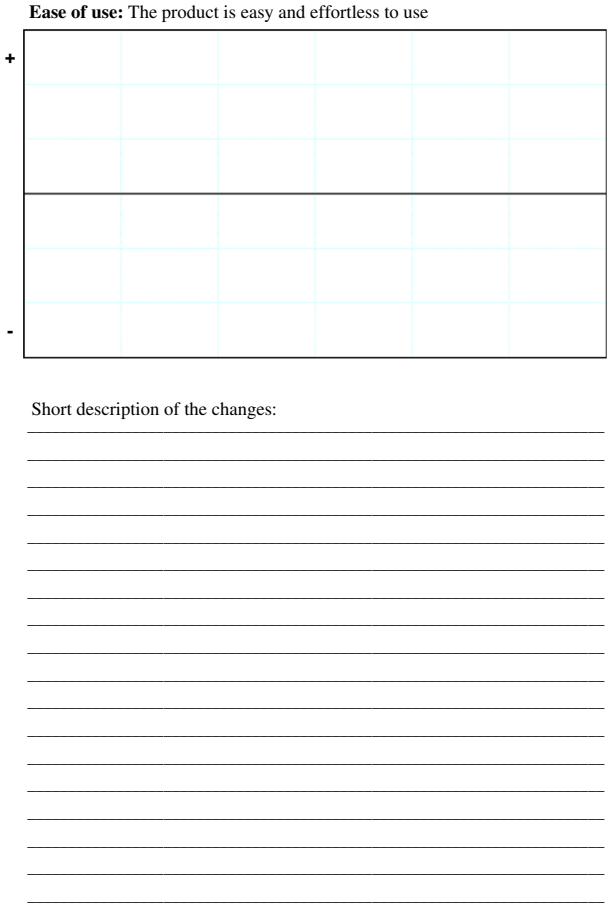

Fig. 1. An example of a UX Curve template.

factors found in the literature (Hassenzahl, 2001; Hassenzahl et al., 2000; Hsee and Hastie, 2006). For example, Mahlke (2008) and Mahlke and Thüring (2007) distinguishes the perception of both non-instrumental qualities and instrumental qualities (utility, usability) as part of user experience. The non-instrumental qualities such as stimulation, identification, innovativeness, esthetics, symbolic and motivational aspects are abstract concepts and difficult to inquire about directly. Thus, we selected one point of view to be general attractiveness (appeal) of the product in the users' own eyes and those of their friends. This concept is close to appeal in Hassenzahl's terms (Hassenzahl, 2001). We expected that this curve would support users to think of more than simply their rational or practical experiences.

In addition, we selected ease of use and utility as two points of view. Ease of use was selected as it is an easier concept for users to understand than usability is. In addition to these three viewpoints of user experience, we wanted to obtain practical information about degree of usage to check if there were long periods of non-usage which might affect user experience trends over time. Degree of usage could be related to quality of experience, at least with voluntarily used consumer products. Thus, if the user experience has an increasing trend, then degree of usage may also increase over time.

We developed five versions of the curve template for gathering data. First, the participants were asked to freely describe their general relationship and user experience by means of the product with the general UX Curve template. The four consecutive curve templates addressed the viewpoints specific to UX: the perceived attractiveness, ease of use, utility, and degree of usage of the product. The first general UX Curve template presented did not have any heading. The other curves were presented with the following headings: Attractiveness: "The product is attractive and interesting in your own eyes and those of your friends." Ease of use: "The product is easy and effortless to use" (as in Fig. 1). Utility: "The product serves an important function for you." Degree of usage: "Degree of usage over time." The degree of usage curve had a slightly different graph area compared to the others, since this aspect obviously cannot be negative: the zero level depicts zero usage.

The method was piloted with 28 researchers and research assistants to determine, first, whether people are able to describe the evolvement of their experiences with a product by drawing a curve, and second, what kinds of curves they draw about the experiences. The participants were asked to select a product they owned and which was important to them. Some of them selected services and one a work-related information system i.e. products that they did not own themselves. The UX Curve method is suitable for these cases as well; we only require that the user has used the product or service for a long time period and not stopped using it after the first negative experiences. Next, the participants were asked to recall the moment when they started to use the product. The first empty template for the general UX Curve was given to them and they were asked to draw a curve describing how their relationship towards the product had changed from the first time they used it until the present. The participants were also asked to describe the reasons for any changes in the area below the graph. When they had drawn the curve and explained the reasons for changes, they were given the next template.

The pilot provided us with confidence regarding the effectiveness of the UX Curve method. The participants elected to describe the user experience of 13 different products and services: these could be categorized as: Facebook (7), mobile phones (6), home and free-time electronics (6), laptops (4), www services (4), and a work related information system (1). All the participants were able to draw the curves and they reported that it was fun apart from writing all the reasons for so many curve types being exhausting. The curve drawings were categorized according to their trends. First, we tested 10 categories to classify all the curve trends, for example a U-form curve, improving negative, improving positive, and changeable. However, the curves were very changeable and for example, a pure U-form was not found. In a second step, we simplified the categorization so as to diminish the divergence of the categories and focused on the curves' three primary trends: improving, deteriorating, or stable. In addition, the participants described the reasons for the changes and large amount of qualitative data was gathered.

4. Validating the UX Curve method

The results of the pilot study showed that users can describe their experiences with the UX Curve method. The common feature of the selected products and services was that they were used on a daily-basis or repeatedly over time. Without repeated use, it would be difficult to draw a curve to describe user experience over time. Next, our goal was to validate the results with a more formal study. The aim was to investigate if the curve trends and their underlying reasons are related to reported user satisfaction, users' willingness to recommend the product and their willingness to buy from the same manufacturer again, i.e. customer loyalty. We used a similar procedure to that of our pilot study but, based on participants' complaints, we attempted to reduce the need for handwritten input by allowing users to think aloud and describe the reasons verbally to the researcher instead of writing them all down. The results obtained from the UX Curve method were then compared with two questionnaires giving information about user satisfaction and customer loyalty.

4.1. Participants and products

In order to keep the curves comparable, we narrowed the selection of products down to just the most popular one in the pilot study: mobile phones that are frequently and consistently used over time. We also limited the usage period to 3–12 months. Karapanos et al. (2009a, 2010) found that 95% of all experiences reported by the iScale related to the first 6 months of use while 75% of them were related to the first month of use, even though the median ownership time was 10 months. Thus, the early experiences seem to be most important, but we wanted to cover a long enough time span to capture the majority of the users' experiences with a mobile phone. The mean usage period was 8.8 months (SD = 2.7). Our goal was to collect a variety of long-term experiences that would allow us to analyze how the results differ for different users and what differentiates the happy and unhappy users. To this end, we selected participants with different backgrounds, ages, genders, mobile phone models, product satisfaction levels, and product usage periods for this exploratory study. All in all, 20 mobile phone users participated in the study. They were recruited through notices in a newspaper and the Nokia customer database and they were all Finns. Their ages ranged from 21 to 53 years (mean 34.7 years, SD = 10.1). The participants' mobile phone models varied, representing two brands; Nokia and LG. 13 out of 20 phones were smartphones. Each participant was rewarded with two movie tickets.

4.2. Procedure

The sessions with the participants consisted of an initial background questionnaire, the curve-drawing session, including the participants' descriptions of the reasons for the changes in their user experiences, and a final questionnaire that enquired into the user's satisfaction and customer loyalty as measured by their willingness to recommend the product to their friends. As the

Please cite this article in press as: Kujala, S., et al. UX Curve: A method for evaluating long-term user experience. Interact. Comput. (2011), doi:10.1016/j.intcom.2011.06.005

participants were Finns, the Finnish language was used in materials and discussion. The mobile phone users participated individually in the sessions, which were of a maximum length of 1.5 h. After the welcoming of the participant and the filling in the initial questionnaire, the curve-drawing session began. First, the researcher asked the participant to recall the moment when they started to use the phone. The researcher handed out the first empty template for the general UX Curve and advised the participant to draw a curve describing how their relationship towards the phone had changed from the first time they had used it until the present. Participants were allowed to tell the reasons to the researcher, but they could also write them down if they wanted to.

After drawing the general curve, each participant was asked to draw four other curves from different UX perspectives: attractiveness, ease of use, utility, and degree of usage. The presentation order of these curves was counterbalanced. We asked the participants to mark the reasons at their approximate locations on the curve. At the end of the evaluation session, the participants filled in a final questionnaire. Finally, the researcher asked participants for feedback on the curve drawing method.

4.3. Materials

4.3.1. Initial questionnaire

The participants' background information was collected through an initial questionnaire. In addition to basic demographic data, we asked questions about their mobile phone usage. For example, we asked about the initial motivation and expectations for choosing and purchasing the phone model they used. We also asked participants for an initial user experience rating for the phone with the question "Is this a good product for you?" To probe whether they had recommended the system to others, we asked "Have you told usage stories or given your opinion about the product to other people? If so, what did you say?"

4.3.2. Curve drawing templates

We collected five different curves with the five curve templates from each participant. The first one was about the participants' general relationship and user experience with the phone. The four consecutive templates addressed specific dimensions of UX: the attractiveness of the phone, its ease of use, its utility, and its degree of usage.

4.3.3. Final questionnaires

At the end of the session, the participants were asked to fill in a questionnaire giving their overall evaluation of the phone, repurchase intentions, and willingness to recommend the phone model to a friend. The first question was the same "Is this a good product for you?" as in the initial questionnaire, since we wanted to see if the curve drawing changed people's overall evaluation of the product. The questionnaire also included questions about meeting expectations: "Did the product meet your expectations?" – repurchase intentions: "How likely it is that you will continue to purchase products from the same manufacturer?" (Reichheld, 2003) and "If your friends were planning to purchase a similar kind of a product, how likely is it that you would recommend this product to them?" (adapted from Reichheld (2003), used also in Chitturi et al. (2008)) – and importance: "How important is this product to you?" We also asked for reasons for their answers, and open-ended questions about the best and worst and any surprising things about the product.

4.4. Data analysis

The curve drawing sessions were audio recorded and transcribed. The reasons explaining the curve trend changes mentioned during the session and written down in the curve template were picked out from the text document and content-analyzed. First, open coding was used to identify themes in the data without ready categories. Second, the identified set of phenomena was categorized using the user experience model by Hassenzahl (2001) and Hassenzahl (2003). The model categorizes user experience issues into pragmatic (utility and usability) and hedonic or non-instrumental. Utility means that the product provides relevant functionality for performing tasks and manipulating environment and usability refers to the ways to access this functionality easily and efficiently (Hassenzahl, 2001). The hedonic categories adapted from Hassenzahl (2003) were stimulation, identification, beauty and evocation. The analysis was performed by one of the authors and then checked by another author.

The first themes identified by open coding fit very well with Hassenzahl's categories (Hassenzahl, 2001, 2003) and the two evaluators were agreeing with the categorization. Only a few reasons (5.7%) needed to be categorized to be 'miscellaneous' as they did not fit the categories and seemed to describe more the change of the user-product relationship over time. These were mostly comparisons with other phones (10 mentions), a user was proud to have better phone than someone else or he was unsatisfied that the phone was not so good than his previous one. Eight users also mentioned that they had become used to the defects. Other miscellaneous reasons were practical problems, expectations, cost issues, and feelings.

The curve forms were categorized according to their trends to improving, deteriorating, or stable. The categorization was

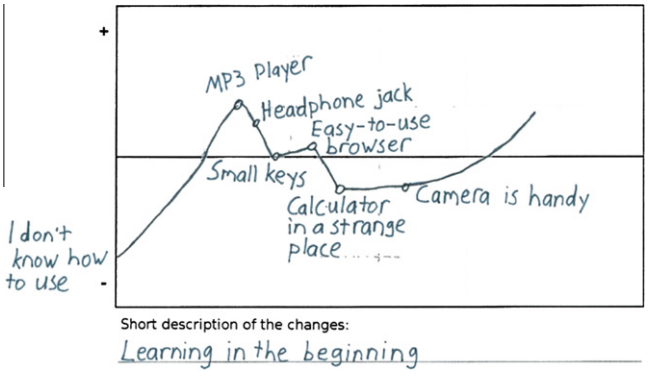

Fig. 2. An example of a general UX Curve drawn by a user, with annotated reasons for changes.

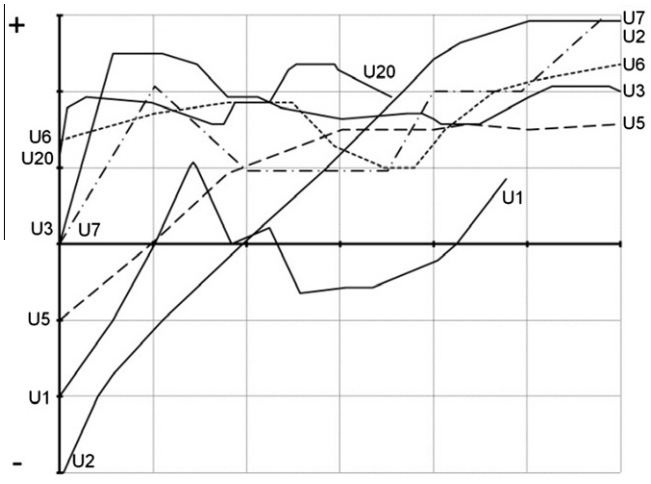

Fig. 3. The improving general UX Curves with user IDs.

performed simply on the basis of whether the starting point of the curve was higher or lower compared to the end point. For example, the curve in Fig. 2 was categorized as being improving as its starting point was lower than its ending point, even though the curve deteriorates in the middle. If the starting and ending points were at the same level, the curve was categorized as stable. As the curves were freehand drawings, they were categorized as stable if there was a very small deviation (less than one millimeter) between the vertical values of the starting and ending points of the curve. However, it can be seen from Figs. 3–10 that the categorization was rather straight-forward to do with the three trend type categories. The relationships between the curve types and the key

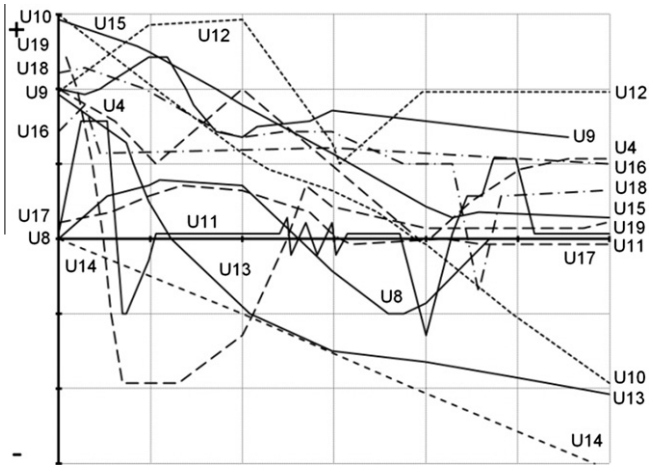

Fig. 4. The deteriorating and stable general UX Curves with user IDs.

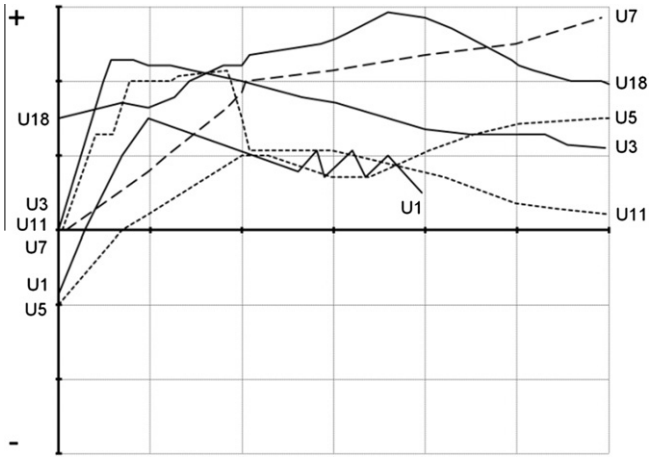

Fig. 5. The improving Attractiveness curves with user IDs.

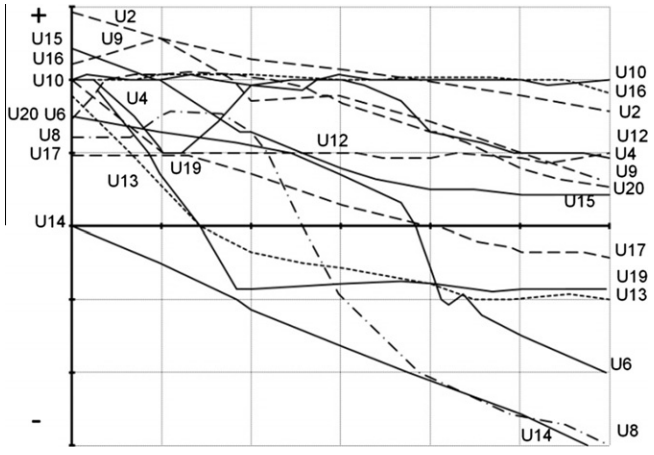

Fig. 6. The deteriorating and stable Attractiveness curves with user IDs.

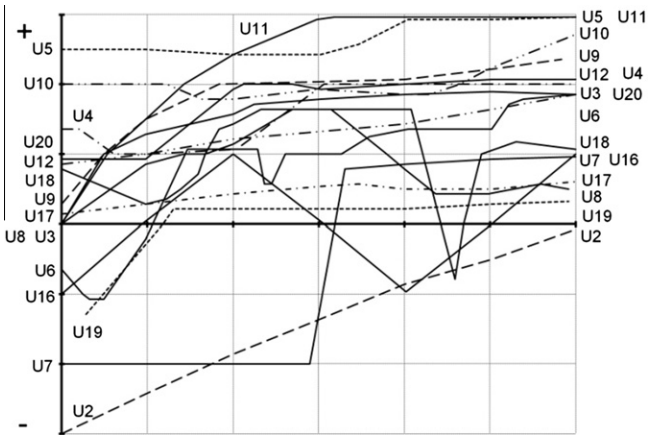

Fig. 7. The improving ease of use curves with user IDs.

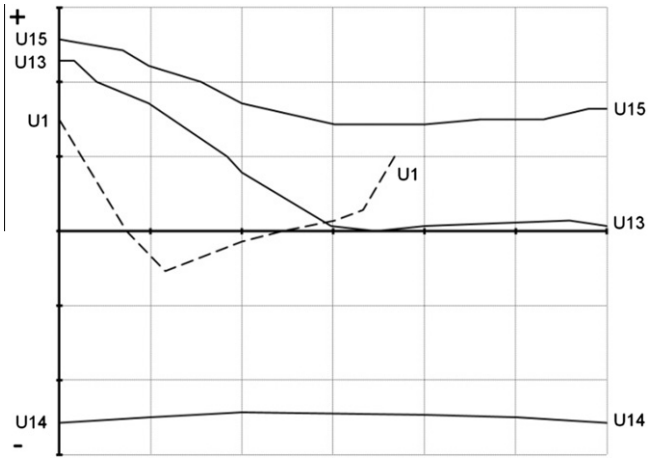

Fig. 8. The deteriorating and stable ease of use curves with user IDs.

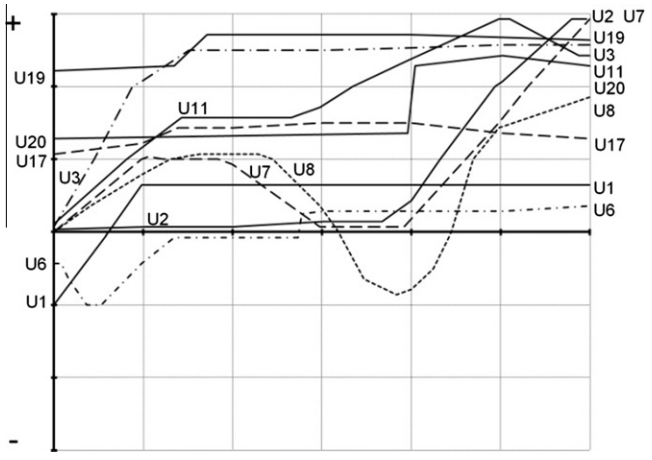

Fig. 9. The improving utility curves with user IDs.

Please cite this article in press as: Kujala, S., et al. UX Curve: A method for evaluating long-term user experience. Interact. Comput. (2011), doi:10.1016/j.intcom.2011.06.005

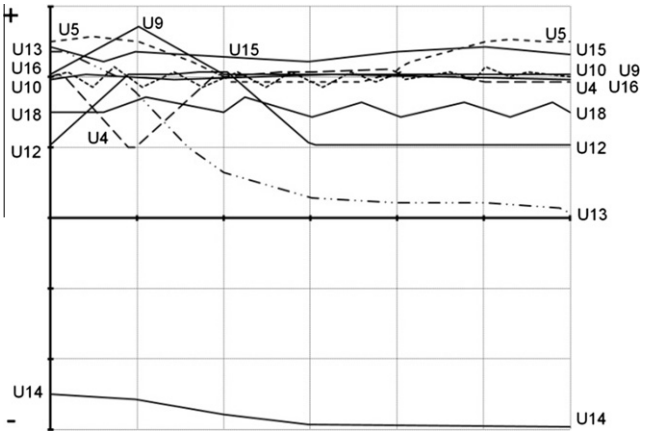

Fig. 10. The deteriorating and stable utility curves with user IDs.

questionnaire results were investigated by statistical analysis. We used the non-parametric Mann–Whitney U test as the size of the sample was small.

5. Results

A total of 100 curves were collected, including 20 general UX Curves, and 20 for each of the four specific UX dimensions. Fig. 2 shows an example of a general UX Curve drawn by a user. The curve shows a typical trend, improving through learning in the beginning. It also shows experiences related to awkward features that have a negative effect on the experience and handy features or pleasurable aspects that improve the experience. The users summarized the reasons that changed their relationship with the phone and then they verbally elaborated upon these experiences. Qualitative content analysis was performed on the verbal data and the trends (improving, deteriorating, or stable) of the curves were analyzed. Finally, statistical analysis was undertaken to determine the relationship between curve trends and overall judgments of satisfaction and willingness to recommend the phone to friends.

5.1. Trends of the curves

Table 1 shows the number of primary trends of the different curve types. Most of the curves were improving, implying an improving experience, but it depended on the particular type of curve: the general UX and Attractiveness curves mostly indicated a decrease and Ease of use, Utility, and Degree of usage mostly showed improvements. Figs. 3–10 show the curves related to General UX, Attractiveness, Ease of use, and Utility to demonstrate how the users drew the curves and how they were categorized into the improving (Figs. 3, 5, 7 and 9) and deteriorating and stable (Figs. 4, 6, 8 and 10) categories. The deteriorating and stable curves were shown in the same figures because there were very few stable curves. The curves, being the participant's stories, seem to reflect their subjective experience when using their phone. For example, it looks as if attractiveness tends to decrease over time, but for some users, perceived attractiveness remained high for very long time periods. Nor do the forms of the curves depend much on the mobile phone model. For example, users U7, U3, U4, and U6 used Nokia E71 phones and their curves represent totally different curve trends.

Table 1
The number of trends of the different curve types.

	Improving	Deteriorating	Stable
General UX	6	12	2
Attractiveness	6	13	1
Ease of use	17	3	0
Utility	10	3	7
Usage volume	13	4	3
Sum	52	35	13

Even though the categorization into improving and deteriorating curves is a very rough analysis of the curve trends, there is a visible difference between the trends. The difference is less dramatic in the Utility curves as users rarely felt that utility had dramatically decreased. On the other hand, the difference between the improving and deteriorating/stable Attractiveness curves seems to be strong. Both improving and deteriorating Attractiveness curves displayed a sharp change right at the beginning of use.

The duration of the usage does not seem to explain the loss of attractiveness. The mean usage time is 8.5 months for improving curves and 8.2 months for deteriorating curves. In addition, the six users who drew curves ending on the negative side of the scale had used their phone on average for 8.5 months. We split the users into those who had been using their phone for 3–7 months ($N = 7$) and those who had been using it for 8–12 months ($N = 13$) and no significant association was found between the groups and their Attractiveness curve trends (Mann–Whitney $U = 33.0$, $p < 0.31$ two-tailed).

5.2. Reasons for the changing user experience

In addition to drawing the curves, the users described the factors that improved their experience over time or caused it to deteriorate. The reasons were also categorized into pragmatic (utility and usability) and hedonic issues based on Hassenzahl's (Hassenzahl, 2003) model of user experience. Table 2 shows the categories and number of reasons related to each category.

When drawing the General UX Curve, users gave an average of 6.5 reasons (3.5 positive and 3.1 negative, $N = 20$). The pragmatic utility reasons were related to functionality, durability, and practicality. For example, they describe how "the programs remain on and that is a bad error as the battery runs down in a day", "the phone became better when I bought a bag for it, now I am not always losing the back cover", "the initial enthusiasm was reduced when I noticed that the quality of the camera was surprisingly low". The hedonic reasons were mostly related to the stimulation, identification and beauty. Four mentions were related to the quality of the phone compared to the user's previous one, two mentions were related to how the phone looks in the eyes of friends, and one user mentioned personalization as a positive factor improving the curve. As Fig. 11 shows, the mean number of positive reasons was higher for those users who drew an improving curve ($N = 7$) than for those users who drew the deteriorating curves ($N = 10$; the stable curves are not included in the figure). Correspondingly, the mean number of negative reasons was lower for those users who drew improving curves compared to those who had deteriorating curves. The quality of reasons related to the improving or deteriorating curves were rather similar, but technical faults and bugs were mentioned more often related to the deteriorating curves than the improving curves (eight vs. two users).

The Attractiveness curves elicited the highest number of reasons, the mean being 7.0 (4.3 positive and 2.7 negative). The reasons mentioned were very similar compared to the general curve, but the focus was more on hedonic reasons. All except one of the users mentioned something related to the design, looks, and esthetics of the phone. For example, they mentioned that "the surface of the phone is shiny and it is very nice looking", "the phone has not started to creak anywhere", and "the phone

Please cite this article in press as: Kujala, S., et al. UX Curve: A method for evaluating long-term user experience. Interact. Comput. (2011), doi:10.1016/j.intcom.2011.06.005

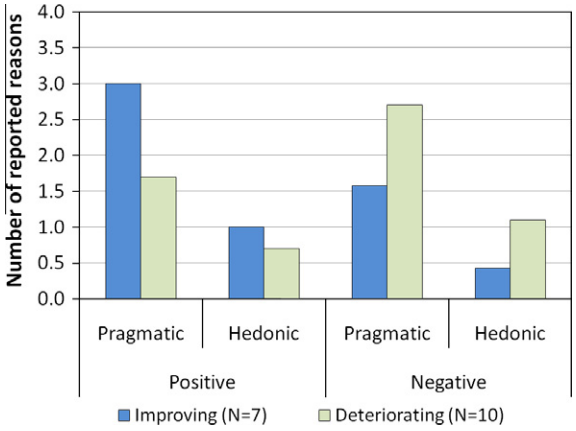

Fig. 11. The mean numbers of reported reasons for improving or deteriorating general UX curves.

feels nice in my hand". Four users mentioned that wear caused the look of the phone to deteriorate over time. An oft-mentioned reason for improvement was related to social status: 13 out of 20 users mentioned how good the phone looked in the eyes of others or how their friend had the same model. The improving Attractiveness curves were again more related to positive reasons than were the decreasing curves (Fig. 12). However, both types of curves had about the same number of negative reasons.

As expected on the basis of von Wilamowitz-Moellendorff et al. (2006), the Attractiveness curves had a tendency to deteriorate as a result of the loss of novelty (eight mentions) and comparison with other phones (eight mentions). In addition, faults and bugs were mentioned five times in relation to the negative curves, but they were not mentioned in relation to the improving curves.

The Ease of use curves elicited 5.6 reasons on average, 3.0 positive and 2.6 negative. The users mentioned reasons such as easy- or difficult-to-use features and general ease of use or familiarity. For example, one of the users stated that "there were confusing terms in the menu: what to expect to find under tools and applications?". They also mentioned reasons for the phone being comfortable or uncomfortable to use, such as the battery life not being long enough, the camera being slow to use, or a good or a missing feature. All except three of the Ease of Use curves were improving and seven of the users mentioned learning as the reason for the ease of use improving. This time, the users with deteriorating curves reported more reasons on average, both positive and negative, than the users with improving curves. The deteriorating curve

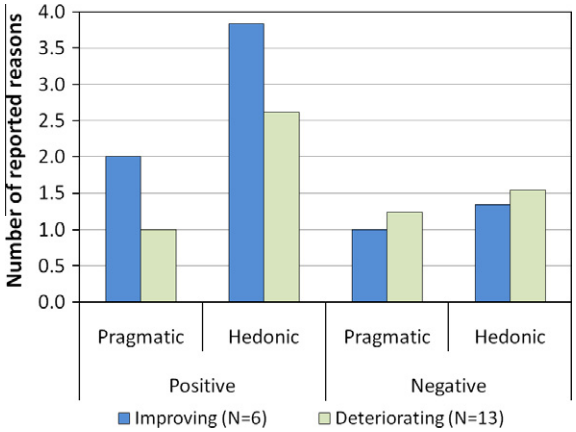

Fig. 12. The mean numbers of reported reasons for improving or deteriorating Attractiveness curves.

experiences related to difficulties caused by technical defects or individual frustrating features.

The Utility curve elicited 4.5 reasons on average, 3.4 positive and 1.1 negative. For example, the users described that "the camera is bad, so I have not had any utility from it" and "Internet is working very well and the phone has Wlan, at home, you do not need to turn on the computer as you can check with the phone". Similar to ease-of-use, most of the users experienced the usefulness as improving over time and they reported that they found new functions and services and learned to utilize them fully over time. The users with deteriorating curves reported more reasons on average, both positive and negative, than the users with improving curves. The deteriorating utility was seen as being caused by technical faults or poor quality that made it impossible to utilize all the functions and applications.

Degree of usage curves were most often increasing for the same reasons as the Utility curves. Using a mobile phone is part of everyday life and a necessity. Only four of the users reduced their use of the phone, two because of changes in their lives, one because of usage problems with the battery, and one because the earphones for the radio stopped working.

5.3. Relation of the curve trends to user satisfaction and recommendation

More positive pragmatic and hedonic reasons, as well as fewer negative pragmatic reasons, were associated with improving curves. This suggests that the improving curves are related to positive UX and user satisfaction. To test this hypothesis, we compared the results of the final questionnaire to curve trends. The improving trend of both the general and Attractiveness curves seemed to be related to user satisfaction and the likelihood of users recommending their phone to their friend. However, only the trends of the Attractiveness curve had a statistically significant relation with user satisfaction, met expectations, and recommendations as shown in Table 3.

Thus, the results suggest that the satisfied users draw an improving trend in the Attractive curve and they report a higher

Table 2
The categories of the reasons for improving or deteriorating user experience.

	Positive	Negative
Utility	114	36
Usability	78	108
Stimulation	21	12
Identification	31	13
Beauty	21	12
Esthetics in interaction	10	7
Evocation	0	1
Miscellaneous	12	6
Sum	287	195

Table 3
The effect of the trend of the Attractiveness curve on user satisfaction and recommendation (scales 1–7, 1 = highly unlikely, 7 = highly likely).

Question	Trend	Mean	*p*
Is this a good product for you?	Improving Deteriorating	6.2 4.7	.022
Do you think that the product has met your expectations?	Improving Deteriorating	6.2 4.1	.017
If your friend was planning to purchase a similar kind of a product, how likely is it that you would recommend this product to him/her?	Improving Deteriorating	6.3 4.3	.022

Please cite this article in press as: Kujala, S., et al. UX Curve: A method for evaluating long-term user experience. Interact. Comput. (2011), doi:10.1016/j.intcom.2011.06.005

likelihood of recommending the products to their friends. The likelihood of recommending was related to the users' actual behavior. Users likely or very likely to recommend their phone had also told positive stories and sometimes mentioned a single negative feature. The users who were not very likely to recommend their phone had told others about the problems with their phone. As mentioned above, the length of the usage period did not explain the differences in attractiveness. Neither did the participants change their satisfaction much as a result of the UX Curve task. In the initial questionnaire, the mean satisfaction was 5.5 and in the final questionnaire it was 5.1. The correlation between the questionnaires was .85, showing a strong level of consistency with the satisfaction question.

We found no statistically significant relationship between attractiveness and repurchase intentions, as many of the disappointed users were still very loyal to the brand. As the users had already had several mobile phones, they based their loyalty on their experiences with previous models as well. For example, one user mentioned that he has had 15 phones, only one had been a brand other than his current one and it would take considerable effort to make him to change brand. In addition, most of the phones were produced by a local company and people are generally very loyal to the brand.

Furthermore, the temporal trend of degree of usage was not related to user satisfaction. Mobile phone use is necessary and its user experience has a smaller impact on degree of usage. (However, degree of usage did not differentiate between the amount of time and the intensity of the use.) The degree of usage curves were used to check that the users really had used the product during the ownership period. The degree of usage curves described many changes related to contextual factors, such as the holiday season, but they did not show any long periods of not using the products.

5.4. Lessons learned from using the UX Curve method

All in all, the UX Curve method was reasonably clear to a wide range of users. All the participants were able to draw the curves, although two of them found drawing odd at first and needed encouragement and help at the beginning. Curve drawing seems to support users in remembering a greater number of details of their use when compared to the CORPUS interviewing (von Wilamowitz-Moellendorff et al., 2006). The number of individual experiences described per user was 28.7 in a 1.5 h session (including two questionnaires) compared to 3.4 experiences in a 45 min session with the CORPUS.

All in all, the users reported 287 positive and 195 negative reasons that affected their user experience. The curve drawing was evaluated to be "slightly funny, but more creative than traditional questionnaires", "interesting and challenging", "at first difficult" and "nicely interactive". Some of the users mentioned that they could not describe the time frame of the events exactly: "it was difficult to split the events to the time line; I did not remember when something happened". They also used different ways of annotating the curves. For example, two of the users marked a positive reason which improved their user experience at the peak of the curve, after which it actually started to deteriorate, while another user marked a negative reason at a similar peak. That is, the exact placing of the experiences were not consistently marked by different users. However, this varied use of the annotating did not affect the current study as the exact timings of the experiences were not analyzed Clearly, the UX Curve is not an exact representation of experiences in the usage period, but rather users' approximate reconstructions of events that are meaningful to them.

The resulting qualitative data is not as rich as in field studies, but it demands fewer resources to analyze, can cover a longer time period, and more users can be involved to improve the generilizability of the results. Compared to semantic differential questionnaires (Hassenzahl, 2001), which provide only numeric results of how interesting, costly, exciting, exclusive etc. the product feels, the UX Curve provides more qualitative data about the reasons that contribute to user experience. The analysis of qualitative data is more demanding than the quantitative data, but for product development purposes a concrete listing of findings and their frequencies among users could be a sufficient level of analysis and most of the findings are directly applicable in design work. The strength of the UX Curve method is that users tell of experiences that are meaningful and memorable and thus help designers identify issues that create positive (or help avoid negative) experiences and affect customer loyalty. On the other hand, as seen from the examples given, some user descriptions such as "the quality of camera was surprisingly low" are not detailed enough to be used in design work and further questioning or other methods such as usability testing may be needed to uncover the details that users are not able to tell.

6. Discussion and conclusions

In this paper, we report our experiences of the UX Curve method in studying how user experience of mobile phones develops over the first 3–12 months of real use. In particular, the goal of the study was to evaluate the usefulness of the method and its relation to user satisfaction and users' recommendation behavior. The results suggest that the UX Curve method is useful for evaluating the quality of long-term user experience as it provided rich qualitative data describing the quality of long-term user experience and identifying the major individual experiences that changed users' attitude towards the product.

In addition to the qualitative data, the UX Curve also provided quantitative data about the trends of user experience over time. The trend of the Attractiveness curve also seemed to be related to user satisfaction and recommending behavior which is found to relate to customer loyalty (Reichheld, 2003). The improving trend of the Attractiveness curve had a statistically significant association with user satisfaction, met expectations, and positive stories and willingness to recommend the product to others.

The UX Curve method relies on users' memories of their experiences; as these are not exact representations of the real experiences and some of these memories are more like retrospective, summative judgments, the method does not uncover all the important details of experiences for design purposes. However, as Norman (2009) and Karapanos et al. (2010) argue, we believe that these memories are essential as they are most meaningful to users. The relation of the Attractiveness curve trend to user satisfaction and recommending behavior also confirmed that the recalled experiences affect behavior and in this way, the UX Curve helps designers to identify the most influential experiences.

The Attractiveness curve seems to be the most powerful curve type, although asking users to draw curves for several UX properties produced more data than one curve alone could. The Attractiveness curves provided the largest number of reasons for explaining the changes in user experience. In addition, they seem to provide the broadest view of the users' experiences by revealing both pragmatic (goal-oriented) and hedonic (pleasure-producing) quality perceptions. While drawing the other curves, users focused more on describing functions and the quality of functions. Moreover, the improving trend of the Attractiveness curves increases the likelihood that the users will recommend the product to others. This result shows that good UX design can have financial effects, as the relationship between customer recommendations and profitable product growth has been shown in earlier research (Reichheld, 2003).

Please cite this article in press as: Kujala, S., et al. UX Curve: A method for evaluating long-term user experience. Interact. Comput. (2011), doi:10.1016/j.intcom.2011.06.005

The importance of attractiveness and hedonic quality perceptions is in line with previous research showing that both pragmatic and hedonic qualities contribute to the overall value of the product (Hassenzahl et al., 2000; Hassenzahl, 2001) and that the relevance of the hedonic aspects seems to increase over time (Fenko et al., 2010). In our study, the hedonic aspects such as pleasure, beauty, and stimulation improved the attractiveness of the mobile phones over time. In particular, as Karapanos et al. (2009b) suggest, supporting self-identity was important to the users. Many of the users mentioned how good the phone looked in the eyes of others and some also described how pleasurable it was that the phone's properties matched their taste or they could personalize it to their taste. As Karapanos et al. (2009b) point out, the focus of research work has been on the product qualities that dominate in early use and our results support their idea that we should study prolonged experiences and understand how a product becomes meaningful in a person's life. The UX Curve is a practical tool for identifying these product qualities that has meaning to users in prolonged use and the reasons which change users' perceptions.

The results suggest that the improving trend of user experience has a relation to user satisfaction as the improving trend of the Attractiveness curves has a statistically significant relation to user satisfaction, met expectations, and recommendations. This observation closely aligns with psychological studies showing that in retrospective evaluations of affective experience, peak and end points influence evaluation the most (Fredrickson, 2000). For example, when comparing pain experiences, people prefer more overall pain if the end of the experience is better (Kahneman et al., 1993). Zauberman et al. (2006) found that hedonic evaluation tasks (rating the feelings generated by a given experience) lead to stronger preferences for improving experiences than do rational informational evaluation tasks. The 'peak-end' bias is assumed to occur as peak and end experiences are well recalled at the time of evaluation (Hsee and Hastie, 2006). The improving trend of a curve means that the highest peak is also often at the end. Thus, design should consider the temporal evolution of user experience and pay more attention to means of improving product attractiveness over time. Our results also resonate with the claims by Norman (2009) and Karapanos et al. (2009a) that memories are more important than single experiences in users' overall evaluations of products. Twelve out of 20 users reported that their phone was a good product for them. However, all the users reported both pleasant and unpleasant experiences with their phones. As satisfied users also reported negative experiences, the results suggest that not even a memorable negative experience can ruin the overall UX.

The results show that user experience is personal – as suggested by many researchers. Different users had diverging reactions even to the same phone model. As user experience depends on the context, user characteristics also have a great role in forming user experience. For example, a technically oriented user may not face the same usability problems than others and this may give rise to different experiences. Thus, a good cross-selection of typical users is needed to give better evaluations of a particular product. In addition, the qualitative data can be analyzed to identify the recurring reasons for improving or deteriorating user experience over time.

We expected perceived attractiveness to correlate with repurchase intentions. However, no statistically significant relationship could be found. It thus seems that repurchase intentions are influenced by a broader set of brand-related aspects. The users surveyed had had several previous mobile phones and a single model's user experience does not necessarily eliminate a person's loyalty to a company. In this study, some users were increasingly attracted to their mobile phone over several months. This suggests that UX can improve over long time periods, contrary to the findings of von Wilamowitz-Moellendorff et al. (2006). For example, good design matches the user's identity as it looks good in the eyes of others, bringing a lasting positive experience. The finding that the improvement in attractiveness in particular is related to the users' willingness to recommend the product to friends supports the idea that hedonism (or pleasurability), rather than pragmatics, drives bonding to a product (Karapanos et al., 2010).

On the basis of the users' most influential individual experiences, we can see that there are competing forces affecting the trend of user experience over time. Overall, we were able to identify the same phenomena as von Wilamowitz-Moellendorff et al. (2006) and Karapanos et al. (2009b). Ease of use improves over time as users become more familiar with a product and learn how to use it. In addition, losing the initial novelty and interest and comparison with other mobile phones are reasons for the deterioration of UX. We also found that technical faults and bugs in particular were related to the deterioration of UX and repeated impracticalities and problems seem to increasingly irritate users over time and cause their UX to deteriorate. For example, one of the users described how after she had made the same error five times while trying to write a text message, she was ready to throw the phone at the wall. She stated that the problem has started to increasingly bother her as time goes on.

The UX Curve method seems to be a reasonably cost-efficient method for studying UX over time. Compared to longitudinal studies, it is a straightforward method that supports users in recalling essential issues related to long-term user experience. Our pilot study suggests that the method is suitable for evaluating various products that are used on a daily basis or repeatedly over time. The UX Curve is best suited for gathering feedback about a product already in the market and supporting the long-term development of customer loyalty with subsequent products or product generations. However, the method may be less suitable for analyzing user experience of a product that is currently in development. Other methods, such as usability tests, are also needed to provide data on more exact user interface details.

The resulting qualitative data are not as rich as in field studies, but it is not so demanding to analyze and can cover a longer time period. In this study, the analysis required a moderate effort, as the qualitative data was content analyzed and categorized. For product development purposes, lighter analysis would probably be sufficient for listing the reasons that reduce or improve attractiveness, finding out which reasons are reported most often, and then trying to fix the negative points and amplify the positive ones. This kind of studies is less onerous to perform, more users can be involved compared to field studies, and the generalizability of the results will be easier to confirm.

Compared to semantic differential questionnaires (Hassenzahl, 2001), the UX Curve method efficiently reveals reasons for improving or deteriorating user experience, which is very useful information when designing for a delightful user experience. Compared to the CORPUS technique, where the researcher draws the curves (von Wilamowitz-Moellendorff et al., 2006), our chronological curve drawing by the users themselves exposed more individual experiences per user. The number of individual experiences described per user was 28.7 in a 1.5 h session (including two questionnaires) compared to 3.4 experiences in a 45 min session with the CORPUS. While CORPUS forces the participant to analytically identify changes starting from the present day and going backwards, our method makes the participant think about their ongoing experiences. Thus, the recall of each experience may further cue more experiences and contextual details and result in richer memories (Karapanos et al., 2009a).

In the current version of the method, a researcher explained to the users what to do and let them explain the reasons for the changing UX verbally. We have also piloted curve drawing as a

Please cite this article in press as: Kujala, S., et al. UX Curve: A method for evaluating long-term user experience. Interact. Comput. (2011), doi:10.1016/j.intcom.2011.06.005

questionnaire with a group of users and it seems to work as well, but writing down the explanations requires more effort from the participant. Tools like iScale by Karapanos et al. (2009a, 2010) support the gathering of curve drawing data online, thus allowing large numbers of people to be surveyed.

The study was exploratory in its nature, as the goal was both to develop a method for helping users to remember how their user experience evolves over time and to investigate the temporal aspects of user experience. The results provide new insights into long-term user experience, but also raise interesting challenges for further research. In our future work, we will extend the UX Curve studies to check the generalizability of the results and the applicability of the method to other product and service categories. We will also work with product development teams to establish the usefulness of the results for product UX design and evaluation purposes.

References

Chitturi, R., Raghunathan, R., Mahajan, V., 2008. Delight by design: the role of hedonic versus utilitarian benefits. Journal of Marketing 72, 48–63.

Courage, C., Jain, J., Rosenbaum, S., 2009. Best practices in longitudinal research. In: Proceedings of the International Conference Extended Abstract on Human Factors in Computing Systems, pp. 4791–4794.

Desmet, P., Hekkert, P., 2007. Framework for product experience. International Journal of Design 1 (1), 57–66.

Fenko, A., Schifferstein, H.N.J., Hekkert, P., 2010. Shifts in sensory dominance between various stages of user-product interactions. Applied Ergonomics 41, 34–40.

Forlizzi, J., Battarbee, K., 2004. Understanding experience in interactive systems. Proceedings of the DIS 04 Conference on Designing Interactive Systems: Processes, Practices, Methods and Techniques, pp. 261–268.

Fredrickson, B.L., 2000. Extracting meaning from past affective experiences: the importance of peaks, ends, and specific emotions. Cognition and Emotion 14 (4), 577–606.

Garrett, J., 2006. Customer loyalty and the elements of user experience. Design management review. Winter 17 (1), 35–39 (ABI/INFORM Global).

Hartmann, J., Sutcliffe, A., Angeli, A., 2008. Towards a theory of user judgment of aesthetics and user interface quality. ACM Transactions on Computer-Human Interaction (TOCHI) 15 (4), 1–30.

Hassenzahl, M., 2001. The effect of perceived hedonic quality on product appealingness. International Journal of Human – Computer Interaction 13 (4), 481–499.

Hassenzahl, M., 2004. The interplay of beauty, goodness, and usability in interactive products. Human-Computer Interaction 19 (4), 319–349.

Hassenzahl, M., Tractinsky, N., 2006. User experience – a research agenda. Behaviour & Information Technology 25 (2), 91–97.

Hassenzahl, M., Platz, A., Burmester, M., Lehner, K., 2000. Hedonic and ergonomic quality aspects determine a software's appeal. In: Turner, T., Szwillus, G. (Eds.), Proceedings of the CHI Conference. ACM, pp. 201–208.

Hassenzahl, M., 2003. The thing and I: understanding the relationship between user and pro Hassenzahl, M. The thing and I: understanding the relationship between user and product. In: Blythe, M., Overbeeke, C., Monk, A.F., Wright, P.C. (Eds.) Funology: From Usability to Enjoyment, pp. 31–42.

Hsee, C.K., Hastie, R., 2006. Decision and experience: why do not we choose what makes us happy? TRENDS in Cognitive Sciences 10 (1), 31–37.

Jordan, P.W., 1998. Human factors for pleasure in product use. Applied Ergonomics 29 (1), 25–33.

Jordan, P.W., 2000. Designing Pleasurable Products: An Introduction to New Human Factors. Taylor & Francis, London.

Kahneman, D., Fredrickson, B.L., Schreiber, C.A., Redelmeier, D.A., 1993. When more pain is preferred to less: adding a better end. Psychological Science 4 (6), 401–405.

Kahneman, D., Krueger, A.B., Schkade, D.A., Schwarz, N., Stone, A.A., 2004. A survey method for characterizing daily life experience. The day reconstruction method. Science 306, 1776–1780.

Karapanos, E., Hassenzahl, M., Martens, J.-B., 2008. User experience over time. In: CHI'08 Extended Abstracts, ACM, pp. 3561–3566.

Karapanos, E., Martens, J.-B., Hassenzahl, M., 2009a. Reconstructing Experiences through Sketching. <http://arxiv.org/abs/0912.5343>.

Karapanos, E., Zimmerman, J., Forlizzi, J, Martens, J-B., 2009b. User experience over time: an initial framework. In: CHI '09 Proceedings, ACM, pp. 729–738.

Karapanos, E., Martens, J.-B., Hassenzahl, M., 2010. On the retrospective assessment of users' experiences over time: memory of actuality? In: CHI'10 Extended Abstracts on Human Factors in Computing Systems, ACM.

Karapanos, E., Zimmerman, J., Forlizzi, J., Martens, J.B., 2010. Measuring the dynamics of remembered experience over time. Interacting with Computers 22 (5), 328–335.

Mahlke, S., 2006. Aesthetic and symbolic qualities as antecedents of overall judgments of interactive products. In: People and Computers XX – Engage, Proceedings of HCI '06, Springer, London.

Mahlke, S., 2008. User Experience of Interaction with Technical Systems. Theories, Methods, Empirical Results, and their Application to the Design of Interactive Systems. Saarbrücken. VDM Verlag, Germany.

Mahlke, S., Thüring, M., 2007. Studying antecedents of emotional experiences in interactive contexts. In Proceedings of the CHI Conference, ACM, pp. 915–918.

Mendosa, V., Novick, D.G., 2005. Usability over time. In: ACM 23rd International Conference on Computer Documentation, ACM, pp. 151–158.

Norman, D.A., 2009 Memory is More Important than Actuality. Interactions, March + April, pp. 24–26.

Oishi, S., Sullivan, H.W., 2006. The predictive value of daily vs. retrospective well-being judgments in relationship stability. Journal of Experimental Social Psychology 43, 460–470.

Reichheld, F.F., 2003. The one number you need to grow. Harvard Business Review 81 (12), 46–54.

Ross, W.T., Simonson, I., 1991. Evaluations of pairs of experiences: a preference for happy endings. Journal of Behavioral Decision Making 4, 273–282.

Schrepp, M., Held, T., Laugwitz, B., 2006. The influence of hedonic quality on the attractiveness of user interfaces of business management software. Interacting with Computers 18, 1055–1069.

Tractinsky, N., Zmiri, D., 2006. Exploring attributes of skins as potential antecedents of emotion in HCI. In: Fishwick, P. (Ed.), Computing. MIT Press, Cambridge.

Van der Heijden, H., 2004. User acceptance of hedonic information systems. MIS Quarterly 28 (4), 695–704.

Vermeeren, A., Lai-Chong Law, E., Roto, V., Obrist, M., Hoonhaut, J., Väänänen-Vainio-Mattila, K., 2010. User experience evaluation methods: current state and development needs. In Proceedings of the NordiCHI Conference.

von Wilamowitz-Moellendorff, M., Hassenzahl, M., Platz, A., 2006. Dynamics of user experience: how the perceived quality of mobile phones changes over time. In: User Experience – Towards a Unified View, Workshop at the 4th Nordic Conference on Human-Computer Interaction, pp. 74–78.

Zauberman, G., Diehl, K., Ariely, D., 2006. Hedonic versus informational evaluations: task dependent preferences for sequences of outcomes. Journal of Behavioral Decision Making 19, 191–211.

Please cite this article in press as: Kujala, S., et al. UX Curve: A method for evaluating long-term user experience. Interact. Comput. (2011), doi:10.1016/j.intcom.2011.06.005