

See discussions, stats, and author profiles for this publication at: <http://www.researchgate.net/publication/262286562>

# Emotions, experiences and usability in real-life mobile phone use

CONFERENCE PAPER · APRIL 2013

DOI: 10.1145/2470654.2466135

---

CITATIONS

4

---

READS

25

## 2 AUTHORS:



Sari Kujala

Aalto University

45 PUBLICATIONS 639 CITATIONS

SEE PROFILE



Miron-Shatz Talya

Ono Academic College

47 PUBLICATIONS 218 CITATIONS

SEE PROFILE

# Emotions, Experiences and Usability in Real-Life Mobile Phone Use

**Sari Kujala**

Aalto University, Department of Design  
P.O. Box 31000  
FI-00076 Aalto, Finland  
sari.kujala@aalto.fi

**Talya Miron-Shatz**

Center for Medical Decision Making,  
Ono Academic College  
104 Zahal St. Kiryat Ono, Israel, 55000  
talyaam@ono.ac.il

## ABSTRACT

Positive emotional experiences with an interactive product are assumed to lead to good user experience and, ultimately, to product success. However, the path from emotional experiences to product evaluation may not be direct, as emotions fluctuate over time, and some experiences are easier to recall than others. In this study, we examined emotions and experience episodes during real-life mobile phone use over a five-month period. The goal is to understand how emotions and memories are related to overall evaluation of a product: usability, user experience and behavioral intentions. The results show that both emotions and how people remember them had strong unique roles in the overall evaluation of the product. Positive emotions were mostly related to good user experience and negative emotions to low usability. In the early stages of use, users overestimated their positive emotions and seemed to focus on user experience, the importance of usability increased over time.

## Author Keywords

User experience; usability; emotions; memories; word of mouth; user satisfaction; day reconstruction method; mobile phone

## ACM Classification Keywords

H.5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous.

## General Terms

Human Factors; Design.

## INTRODUCTION

Positive user experience is an important goal in design. Companies consider good user experience vital to continuous commercial success and it is believed to improve customer loyalty. For example, Jordan [17] suggests that if people have pleasurable experiences with a product, they are more willing to buy the next product from the same company. A good user experience may also lead

users to recommend the product to their friends, which brings in new customers [cf. 33].

It is still unclear what user experience exactly is [6, 25], but users' emotional responses to interactive products are frequently viewed as essential in user experience [6, 10, 16, 26] and emotion appears to help users evaluate outcomes when interacting with products [6]. For example, according to Mahlke and Thüring's [27] model, user experience consists of emotional reactions and perceptions of instrumental and non-instrumental qualities.

In this paper, we discuss the role of emotions in real-life mobile phone use over a five-month period. As emotions fluctuate over time, it is necessary to know how users remember them and to define their role in usability, user experience and ultimately in product success.

The role of emotions over extended time periods is not well understood. In studies to date, the time periods examined are relatively short [24]. For example, Mahlke and Thüring [27] found empirical evidence that perception of both instrumental and non-instrumental quality influences emotional reactions and the appraisal of interactive systems. However, their experiment lasted an average of only 60 minutes. Similarly, Tuch et al. [39] found that the visual complexity of websites affects users' emotional reactions, but the experiment included only visual search tasks and it lasted only 90 minutes.

In real life, the relationship between a user and a product develops over months or even years. Users have a substantial amount of emotional experiences with the product – both positive and negative – and emotions may fluctuate from minute to minute and from day to day. When forming an overall evaluation on a product, people do not merely tap into the emotional experience at the moment. Rather, they base their overall evaluation on their entire experience throughout usage, as stored in their memory.

Psychological studies of autobiographical memory have shown that people do not passively undergo emotional experiences, but rather actively interpret the meaning of experiences and construct memories from them (see [12] for a review). Similarly, user experience is frequently conceptualized as constructive [35] and sense making [42].

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

CHI 2013, April 27–May 2, 2013, Paris, France.

Copyright © 2013 ACM 978-1-4503-1899-0/13/04...\$15.00.

The optimal timeframe for evaluating user experience is unclear. Many researchers and practitioners agree that user experience must be assessed during the actual interaction with a product rather than afterwards [25]. However, in practice, such instant evaluations may be made during very short evaluation periods of engagement with the product and may be inconsistent with users' overall interpretations of their experiences as these evolve over time.

As Hassenzahl and Ullrich [11] state, when users are evaluating a product, they make judgments about the overall quality of a product based on their memory for momentary experience episodes. These summary judgments guide users' decisions regarding products: whether they are satisfied with the product and willing to continue using it and even recommend the product to others. The outcome of these decisions can be taken to reflect the success of a product, but how well do users remember their emotional experiences? When users form an overall evaluation of the product, are all experience episodes equally important? Is it the experience episodes themselves or memories of those episodes that determine product success?

In this paper, we present results on how people evaluate their emotions during the first 5 days of mobile phone use, after 2.5 months and after 5 months of use. The study makes a unique contribution in showing that both positive and negative emotions and how people remember them have a strong unique roles in the overall evaluation of the product.

### Memory and experiences

Psychological studies have long shown that human memories are vulnerable to retrospective biases [34]: on-line evaluations of emotional experiences differ from the retrospective overall evaluation. For example, people often base their overall evaluations on the peak and final intensity of the experience presumably because they are well recalled at the time of evaluation [15]. In addition, people tend to overestimate the intensity of the emotional experience compared to their feelings at the time of the experience itself [29, 30]. The discrepancy between the experienced emotions and their evaluation has been demonstrated in the context of experiences ranging from vacations [30, 41] and episodes of pain [38] to everyday events [28, 29].

Also in the field of human-computer interaction, Norman [31] and Karapanos et al. [22] have argued for the importance of memories of experiences (rather than actual experiences) to product success, as it is the memories that will be reported to others and guide future user behavior. Hassenzahl and Sandweg [9] showed that memory influences assessment of perceived usability. In their experiment, these summary assessments did not reflect a whole experience episode, but rather the last task of the whole two-hour test session. They explained that more recent details come to mind more easily.

In addition to memory biases, forecasts of emotional experiences also tend to be more extreme than actual experiences as they focus on core attributes of the activity at the expense of other information [36]. For example, people believe that vacations are highly enjoyable and this belief shapes their predictions and global evaluations of the holidays, even when the actual experience turns out to be disappointing [30]. People forget the negative experiences and base their overall evaluations on expectations and positive experiences [30]. It is psychologically adaptive to view past life experiences in a positive light as this allows us to maintain a positive self-image and to forge positive social relationships [12]. It remains to be examined whether the same happens with commercial products that can be discarded and replaced.

Although global reports are biased and only memories of real experiences, they are important because they seem to be used in predicting behavioral choices [30, 41]. For example, Wirtz's et al. [41] studied students before, during, and after their spring-break vacations and found that the predicted and remembered experiences were more intense than the actual experiences. Recalled memories of a vacation, but not the actual experience, predicted the students' willingness to repeat the experience.

Apart from Hassenzahl and Sandweg's [9] short-term study, there is paucity of scientific evidence supporting the importance of memories in user experience, overall product evaluation and behavioral intentions (e.g. willingness to repeat the experience). Rather, as mentioned above, many researchers highlight the importance of experiences during the interaction [25].

### Capturing emotional experiences over time

In human-computer interaction, studies that track users over several days or weeks are rare because of the associated expense and participant fatigue involved [21]. In particular, it is challenging to recruit and motivate users to participate in studies in which they need to report their experiences daily. To illustrate, Karapanos et al. [21] applied the Day Reconstruction Method (DRM) and evaluated users' daily experience episodes with iPhones for 5 weeks, but they only had six technically-oriented participants.

The Experience Sampling Method (ESM) as reviewed by Csikszentmihalyi and Larson [4] collects information on people's feelings several times a day, and is similar to Ecological Momentary Assessment, applied by Stone et al. [37] which also measures physical sensations. The DRM was developed by Kahneman et al. [18] as a more practical version of ESM for wellbeing research. The DRM reduces participants' burden by imposing a chronological process in the reconstruction of experiences, only once a day. The results are very close to those elicited through ESM [18], yet without interrupting respondents' daily activities.

Karapanos et al. [21] asked their study participants to pick the three most impactful experience episodes of the day with the new phone, either satisfying or dissatisfying, write them down and describe the situations, their feelings and momentary perceptions of the product. They found that many different kinds of experience episodes may take place with the iPhone during the same day, with a distribution that changes over time, going from an orienting learning phase to a final emotional attachment phase, as weeks progress.

**Study overview and hypotheses**

In this study, we examine emotions during real-use mobile phone usage. The goal is to understand how emotions and memories are related to overall evaluation of a product: usability, user experience and behavioral intentions.

Based on the literature review, the overall hypothesis is that users do not have accurate memory of all the details of their experience episodes and related emotions, but rather, they actively interpret the meaning of experience episodes and construct memories of them. It is expected that both the memories and the real emotional experiences are important in the overall evaluation of the product.

We formed three specific hypotheses:

Hypothesis 1: When users recall their emotions related to product use they overestimate the intensity of the emotions compared to the actual intensity during the usage.

Hypothesis 2: Positive emotions and their memories are related to good overall evaluation of the product.

Hypothesis 3: Memories of positive and negative peak use experience episodes are related to the overall evaluation of the product.

**METHOD**

**Participants and products**

Employees of a large university in Finland were recruited to participate in the study. The potential participants were contacted when they ordered a new mobile phone from a university phone service. The users selected their mobile phone for work purposes and the university paid for the phones. All in all, 82 users were sent an invitation e-mail, 42 of them agreed to participate and filled in the initial questionnaire. As the users agreed to participate before they received the phone, some users unexpectedly had sick leaves, holidays or work pressure that prevented them from participating in the study or from completing it. Thus, 24 users filled in the questionnaires of the first six days of the mobile phone use. Of these, 22 users filled follow-up questionnaires after 2.5 months and 5 months.

As our focus was the path from experience episode to memory to evaluation, over an extended period of time, only the responses of these 22 users were included in the analysis. Each participant received the choice of movie tickets or a memory stick as a reward.

The mean age of the participants was 39.1 years (SD=12.2), 12 of them were females and 10 males. They represented different positions (e.g. secretary, laboratory engineer, researcher, professor) and fields (e.g. traffic engineering, design, informatics). The mobile phones that the participants had selected for themselves represented three different brands and five different models (e.g. Nokia E7, HTC Wildfire, iPhone 4S). The participants had not had prior experience of their new mobile phones; one user had used a previous model of the phone.

**Procedure**

The goal was to investigate emotions in real-use contexts and therefore compared to previous studies the study time span was extended to 5 months of usage. In order to ensure good reliability and ecological validity participant burden was reduced by selecting carefully the data collection points to represent time before usage, short-term usage, middle-term usage (2.5 months) and long-term usage (5 months).

Participants were surveyed at several time points using an online tool, Webropol (Table 1). First, 1-3 days before the participants received the new mobile phone, they provided background information about themselves and the phone. Second, the participants reconstructed their experience episodes with the phone and the related emotions during each of the first five days of use using DRM. Third, on the sixth day, the participants rated their overall evaluation of the product and overall satisfaction on an evaluation questionnaire. The participants filled in the evaluation questionnaire again after 2.5 months and 5 months of usage.

**Materials**

*Initial questionnaire*

The participants' background information including prior experience was collected through the initial questionnaire. In addition to basic demographic data, we asked the participants to rate how enjoyable they expected the product to be and how satisfied they expected to be with it.

**Table 1. Measurement points during the five months of use.**

Measurement point	Information gathered
1. Before usage	<i>Initial questionnaire:</i> Background information, expectations
2. During the first five days	<i>Daily DRM questionnaires:</i> Daily use experience episodes and related emotions
3. On sixth day	<i>Evaluation questionnaire:</i> Overall evaluation of the product and recalled emotions, the most memorable peak and low experience episodes
4. After 2.5 months	
5. After 5 months	

The expected enjoyment was rated on a scale ranging from 1 (disagree) to 5 (agree). The statements (adapted from [30, 41]) were “I expect to enjoy using the new product”, “I think using the product will be fun”, and “I will be satisfied with the product”, “I think using the product will be rewarding”, “I feel good about this product”.

#### *Day Reconstruction Method questionnaire (DRM)*

The participants reported daily their experience episodes with their mobile phones and the related emotions during the first five days of usage. The participants followed the Day Reconstruction Method (DRM) protocol by Kahneman et al. [19], which we modified to product use context.

The participants reconstructed the main episodes of the previous day, from when they woke up to when they went to sleep. They then recalled all the experience episodes related to the product during the previous day (either related to the day episodes or not) and to describe their episodes. We defined episodes as product-related experiences, when the participants think about the product, look at it, keep it at hand, use it, or somebody else is looking at it.

First, the participants were asked to name and describe the product related experience episode. Next, the participants were asked to rate the extent to which they experienced various emotions, from 1 (not at all) to 5 (very much) during the episode. As recommended by Kahneman et al. [19], the affective descriptors that were relevant to the usage situations were selected by adopting the DRM’s original affective descriptions. The questionnaire consisted of four descriptors of negative emotions: ‘frustrated’, ‘irritated’, ‘angry’ and ‘tense’. The four positive emotion descriptions were ‘excited’, ‘satisfied’, ‘proud’, and ‘glad’. These measures were averaged to create index variables for positive affect and negative affect. In addition, the participants had the possibility to name and rate one affect by themselves.

After the participants described an experience episode, they were asked if they had more product-related episodes. If the answer was positive, they were asked to describe the additional episodes and to rate the related emotions. The questionnaire was filled in only for those days the participants had experience episodes with their mobile phone, on average for 4.5 days out of the five days investigated.

#### *Evaluation questionnaire*

After the five days, 2.5 months and 5 months usage, the participants were asked to fill in a questionnaire giving their overall evaluation of the phone. This questionnaire included five measures: 1) emotional reactions, 2) behavioral intentions, 3) perceived usability, 4) user experience, and 5) the most memorable experience episodes.

1) *Emotional reactions* were assessed with the same rating questions that were used in the Day Reconstructions Method questionnaire. In addition, the participants were asked whether they had experienced an exceptionally memorable wonderful episode with the mobile phone and then they were asked to rate their emotions during the episode. After that they were asked whether they had experienced an exceptionally memorable awful episode with the phone and they were asked to rate their emotions during it [adapted from 28].

2) *Behavioral intentions* were measured by product success related questions about the participants’ reactions related to the mobile phone. The participants were asked if they would repeat the experience of using the product again and repurchase it using the operationalization of Wirtz et al. [41]. The participants were asked: ‘Would you start using this same product over again (assuming that you can start again and you know what you now know)?’ Their willingness to recommend the product was asked with the question: ‘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 [33], used also in [3]). They were also asked if the mobile phone met their expectations (adapted from [24]) and how willing they are continuing using the same mobile phone.

3) *Subjective usability* was measured with the Usability Metric for User Experience (UMUX) developed by Finstad [5]. It is a four-item Likert scale aimed to measure the three dimensions of usability defined by the ISO 9241-11 standard: effectiveness, efficiency, and satisfaction. The UMUX scale is a shortened version of the ten-item System Usability Scale and according to Finstad’s [5] comparison the measures correlate at a rate 0.80.

4) *User experience* was a measure of hedonic, pleasure producing product quality. It was measured with the AttrakDiff questionnaire 2 [1, 7] that is one of the most used user experience questionnaires [2]. AttrakDiff uses semantic differential in which a user is asked to rate the product using opposite word pairs with a 7-point scale. The questionnaire measures the perceptions of pragmatic and hedonic quality. The pragmatic quality questions are focusing on ease of use and they were excluded in order to avoid overlapping with the subjective usability measure and in order to avoid participant fatigue. Hedonic quality questions focus on stimulation that is related to the novelty and originality of the product, and identification that is the product’s ability to address the need to express one’s self through objects one owns [7]. In addition, the questionnaire includes questions related to the attractiveness of the product (e.g. pleasantness and beauty). The total number of the used AttrakDiff word pairs was 21.

5) *The most memorable experience episodes* were identified by asking the participants about peak and low experience episodes related to the mobile phone. First, they were asked whether they had had an unusually wonderful episode with the phone and then, whether they had had an unusually awful or unpleasant episode with it. Those participants that had had peak episodes were then asked to rate the related emotions similarly to the DRM questionnaire.

### Analysis

The analysis started by establishing the first four summary measures mentioned below and assessing their reliability. In the first measurement after five days of use, Cronbach's alpha coefficient was .94 for behavioral intentions, .74 for subjective usability and .91 for user experience demonstrating high degree of internal reliability of the scales.

For positive emotions Cronbach's alpha was also very high .92. For negative emotions Cronbach's alpha was .78 without Tense and with the 'tense' it was .59. 'Tense' did not correlate negatively with the positive measures as other negative measures, implying that tension is more of a measure of arousal, and not altogether a negative emotion. Thus, the 'tense' measure was excluded from the analysis and we averaged the other three negative measures to create an index variable for negative emotions.

In this study, usability and user experience were measured separately using existing measures although the concepts are partly overlapping. The user satisfaction component of usability is related to user experience. However, the correlations between the used subjective usability and user experience measures were very low and not significant:  $r = -.161$  ( $p < .475$ ),  $.271$  ( $p < .22$ ) and  $.460$  ( $p < .031$ ) after 5 days, 2.5 and 5 months. Also Hornbæk and Law [13] found in their meta-analysis of 73 studies that the effectiveness, efficiency and user satisfaction are correlated although average correlations were rather low, ranging from .247 between effectiveness and efficiency to .164 and .194 between these measures and user satisfaction.

A repeated measure ANOVA was used to compare the means of emotion ratings in different measurement points. Bivariate correlations ( $r$ ) and point-biserial correlations ( $r_{pb}$ ) were used to compute the correlation coefficients among the study variables. A linear mixed model analysis (also called multilevel model) [14] was used to analyze the time effects. Analysis was conducted for behavioral intentions and positive emotions as separate dependent variables in the mixed model. Subject ID was entered as a subject, with time and subjective usability or user experience as fixed effects.

The most memorable peak and low experience episodes were analyzed using open coding. First, open coding was used to iteratively construct the content categories. Some experience episode narratives included two or three different points or sub-episodes so these narratives were

divided accordingly for the content analysis. Six content categories were originally created by one researcher but iterated in collaboration between two researchers and the two researchers independently categorized all the narratives. The inter-rater agreement was substantial, Kappa = .68 with  $p < .001$ .

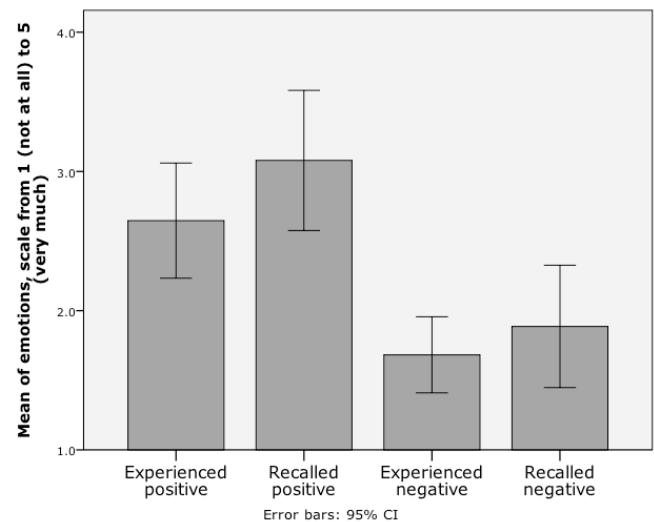
## RESULTS

### Hypothesis 1

In order to assess the prediction that users would overestimate the intensity of their positive and negative emotions after usage, we examined the intensity of emotions during the first five days of usage and after the first five days. As the participants reported all their individual experience episodes and the related emotions for the first five days and they made the overall evaluation of their emotions on sixth day, we could calculate the mean of daily emotions related to the product and then compare the mean to the overall evaluation of the emotions.

As shown in Figure 1 the results confirmed the hypothesis. The users overestimated both their positive and negative emotions. The means of daily experienced positive and negative emotions were 2.65 (SD=.93) and 1.68 (SD=.62) respectively, but the means of overall recalled evaluations were 3.08 (SD=1.13) and 1.89 (SD=.99). The difference is statistically significant for the positive emotions ( $F_{1,21}=9.05$ ,  $p < .007$ ) but not for the negative emotions ( $F_{1,21}=2.74$ ,  $p < .11$ ).

Thus, on the sixth day, when users made their overall evaluation of their new mobile phone, they overestimated particularly their recalled positive emotions compared to the daily experienced emotions.



**Figure 1. Means of daily experienced emotions and recalled overall evaluation of emotions after five days of use.**

**Hypothesis 2**

Hypothesis 2 predicted that positive emotions and their memories will be related to good overall evaluation of the product. This hypothesis was tested by calculating the correlations between the experienced and recalled emotions and product evaluation measures as shown in Table 2.

The results confirm the hypothesis by showing that both the experienced and recalled positive emotions were statistically correlated to good user experience and the product-related behavioral intentions. Also expected enjoyment was related to user experience and the behavioral intentions. The negative emotions were negatively correlated with usability and only the recalled memories correlated with the behavioral intentions related to the product.

Table 3 shows that the mean intensity of the recalled emotions did not change from one measurement to another. However, the role of emotions changed over time. When usability was the dependent variable and time and positive emotions as fixed effects in the mixed model, positive emotions had a significant effect ( $F_{1,37}=21.15, p<.001$ ) so did time ( $F_{2,20}=13.19, p<.001$ ). In addition, there was a significant interaction between time and positive emotions ( $F_{2,20}=10.80, p<.001$ ) showing that the relationship of usability and positive emotions became stronger over time.

Figure 3 shows the change of correlations of usability and user experience with the behavioral intentions over time. When the behavioral intentions measurement was used as the dependent variable, usability had a significant effect ( $F_{1,45}=42.52, p<.001$ ) so did time ( $F_{1,22}=3.61, p<.044$ ). In addition, there was a significant interaction between time and usability ( $F_{1,22}=3.58, p<.045$ ) and the estimates show that the relationship between usability and the behavioral intentions became stronger: the impact of usability on behavioral intentions is increasing over time.

When the behavioral intentions measurement was used as the dependent variable and time and user experience as fixed effects, user experience had a significant effect on behavioral intentions ( $F_{1,46}=7.01, p<.011$ ), but there were no significant time effects ( $F_{2,18}=1.14, p<.34$ ) or interaction between time and user experience ( $F_{2,18}=1.48, p<.25$ ). Thus, user experience is predicting behavioral intentions but the impact of user experience is constant over time.

**Hypothesis 3**

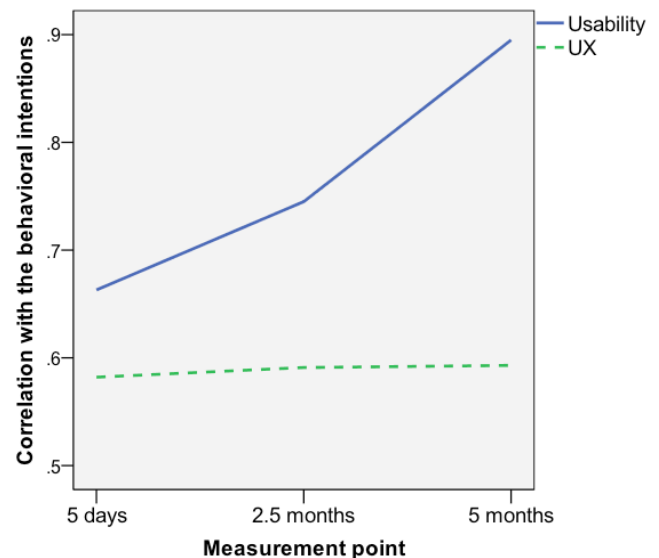
Hypothesis 3 predicted that the memories of positive peak experience episodes and negative low episodes will be related to the overall product evaluation. Five users did not report peak or low experience episodes, but the other 17 users reported altogether 23 positive peak experience episodes (included 27 sub-experiences) and 21 negative low episodes (included 22 sub-episodes). Seven users reported both positive and negative experience episodes, seven users reported only positive episodes and three users reported only negative episodes.

**Table 2. The correlations between the means of daily experienced and recalled emotions after five days of use and the product evaluation measures.**

	Usability	User experience	Behavioral intentions
<b>Positive experienced</b>	.328 p<.145	<b>.736**</b> p<.001	<b>.700**</b> p<.001
<b>Negative experienced</b>	<b>-.429*</b> p<.046	-.093 p<.680	-.315 p<.153
<b>Positive recalled</b>	.374 p<.087	<b>.726**</b> p<.001	<b>.829**</b> p<.001
<b>Negative recalled</b>	<b>-.601**</b> p<.003	-.211 p<.347	<b>-.468*</b> p<.028
<b>Expected enjoyment</b>	.349 p<.111	<b>.839**</b> P<.000	<b>.655**</b> p<.001

**Table 3. Means (and standard deviations) of the product evaluation measures, behavioral intentions and emotions in different measurement points.**

	5 days	2.5 months	5 months
<b>Positive emotions</b>	3.08 (1.13)	3.20 (1.04)	3.15 (1.00)
<b>Negative emotions</b>	1.89 (.99)	1.78 (.68)	1.89 (.77)
<b>Usability</b>	72.2 (20.79)	64.6 (26.28)	67.8 (25.20)
<b>User experience</b>	4.97 (1.16)	4.91 (1.14)	4.86 (1.10)
<b>Behavioral intentions</b>	5.33 (1.33)	5.17 (1.33)	5.18 (1.61)



**Figure 3. Correlation of usability and user experience with the behavioral intentions.**

Positive peak and negative low experience episodes raised strong emotions. Positive peak experience episodes were related to high positive emotion ratings ( $M=4.38$ ,  $SD=.56$ ) and negative low episodes were related to high negative emotions ratings ( $M=4.15$ ,  $SD=.44$ ). Based on the reports of the first five days, the participants used the phone several times a day almost daily. Thus, the peak experience episodes had a strong effect on the overall evaluation of the emotions among the hundreds of ordinary experience episodes during the five months of usage.

The number of peak and low experience episodes were too small to make reliable conclusions of their effects, but point-biserial correlations suggest that peak and low experience episodes are related to overall emotions after five months. Reporting positive peak episodes correlated significantly with the overall positive emotions after five months ( $r_{pb}=.597^{**}$ ,  $p<.003$ ), suggesting that those users who had positive peak episodes also reported more positive overall emotions. Reporting negative low episodes correlated significantly with overall negative emotions after five months ( $r_{pb}=.733^{**}$ ,  $p<.001$ ) suggesting that those users who had negative low episodes also reported more negative overall emotions than rest of the users.

The content analysis of the experience episodes showed that the negative low episodes were mostly related to ease of use (73%). Table 4 provides some examples of the negative low experience narratives. In 16 out of 22 negative episodes, the users described situations where the phone was difficult, uncomfortable or frustrating to use. In three cases, the users reported that the phone was not working properly e.g. after five months one user told that internet and mobile connection were suddenly disappearing and one user told that the charger is working unreliably. In addition, in two cases users described how they felt stupid when using the phone and how some functions were irritating.

Table 5 includes examples of the positive peak experience episode narratives. 14 out of 27 experiences (52%) were related to utility, the users mostly told about a good functionality that enabled them to achieve something useful. Three users also brought up the easiness of doing useful tasks.

Five positive experiences were more experiential relating to sociability and novelty: ‘playing together with my 1.5 old son’ and ‘in the beginning when I started to use the phone, I was really enthusiastic and used lot of time for getting to know the phone’. In addition, two experiences were not directly related to the phone, one person received a call from a friend that she had not seen for a year and one person got an email that his long iterated article was accepted to a good journal.

The content analyses of experience episode narratives also revealed that efficiency and effectiveness seem to be intertwined with user experience or hedonic product quality.

**Table 4. Examples of the negative low experience episodes.**

Description of the low negative experience
<i>The first day! I couldn't even put the PIN code without help although it is usually the easiest task with the new mobile phone. Implementation was followed a general peeping in our room as my neighbor and I were doing settings. Uh! In addition, all in all I felt very stupid (User 8 after five days).</i>
<i>I don't know how to end a call (User 16 after 2.5 months)</i>
<i>Simple things like changing the ringing tone were not easy. It is also somewhat irritating to find numbers with the phone and end the call after a call (User 21 after 2.5 months).</i>
<i>I can't save the phone number of the persons who called, but instead I call them when I try to save the number (User 16 after 5 months).</i>

**Table 5. Examples of the positive peak experiences.**

Description of the positive peak experience
<i>I was managed to put on an alarm for a leaving time and the phone buzzed on the phone so that other people noticed it. I received a comment, hey, you are already a professional in using the phone (User 17 after 5 days.)</i>
<i>The day when I made the phone personal, the day was about one month ago. I updated the operating system and at the same time several Ovi Store applications, mostly newspapers. I was great that I was able to make the local newspaper visible and even for free for the first couple of months (User 3 after 2.5 months).</i>
<i>I used the map and location data when I was lost (User 9 after 2.5 months).</i>
<i>During tram journey I was able to handle several duties by replying to e-mails. I felt efficient in using mobile phone (User 2, after 5 months).</i>

For example, many experiences had social consequences (positive attention or embarrassment) or effects on how the users identified themselves (proud of accomplishment or feeling stupid, irritated or frustrated).

**CONCLUSIONS**

In this study, we investigated emotions during real-life mobile phone use over five months. The goal was to understand how emotions and memories are related to overall evaluation of a product: usability, user experience and behavioral intentions.

**Emotions, memories and overall evaluation**

The results show that both experienced emotions and memories thereof play an important role in overall product evaluation. After five days, users overestimated their



overall positive emotions compared to the daily experienced ones. Both recalled overall positive emotions and experienced ones were related to user experience and behavioral intentions. Thus, it appears that the positive emotions and memories are a strong predictor of product success as measured by the behavioral intentions.

Usability, on the other hand, was related to negative emotions and the memories of the negative emotions were significantly related to behavioral intentions. Over time, good usability was associated with positive emotions parallel with user experience and its effect on behavioral intentions increased while the effect of user experience remained stable over time.

The peak and low experience episodes that the participants reported were presumably the most memorable episodes that they had with the phones. Similarly to emotion ratings, the low experience episodes were related to negative emotions and mainly caused by low ease of use.

Positive peak experience episodes were related to positive emotions and the utility of the phone, e.g. the functionality enabled the person to achieve something positive. The number of peak and low episodes was too small to make reliable statistical tests of their relationship to memories of emotions, but the correlations suggest that the users based on their overall evaluations of the memories on the low and peak experience episodes.

Both the overestimated emotions after five days and the importance of peak and low experience episodes for the overall evaluation support the overall hypothesis that users actively interpret the meaning of their experiences and construct memories of them. Based on this study, these memories of the experience episodes and the related emotions seem to have at least as strong effect on the overall evaluation of the product as the actual experience episodes have.

#### **Emotions, usability and user experience over time**

In the early stages of use, users overestimated their positive emotions and seemed to focus on positive emotions more than on negative emotions.

The finding that positive emotions are related to user experience whereas negative emotions are related to usability clarifies the nature of the usability and user experience concepts, and suggest that good user experience means that the product is emotionally rewarding. Further, good user experience brings positive emotions because it helps to fulfill personal needs as the nature of peak experience episodes suggests. Similarly, Hassenzahl et al. [8] showed this in their analysis of 500 positive experiences with interactive products.

Over time, usability started to bring about more positive emotions. The peak experience episodes also showed the intertwined nature of usability and user experience. Effectiveness was needed in order to achieve the goals of

users but the positive consequences were often described to relate positive user experience, for example feeling effective or getting positive social feedback. Moreover, good usability is rewarding as such, particularly over time as it supports positive self-image.

Based on earlier research results we could have expected that usability is particularly important in the beginning of use for learnability reasons [20, 21]. In Karapanos' et al. four-week study [21] there was a sharp decrease of learnability flaws after the first week of use suggesting the importance of an early learning phase. In another study of Karapanos' et al. [20] pragmatic aspects were relevant in the first measurement during the first week of use but not in the second measurement after four weeks of use.

The low experience episode examples show that this was not the case in the current study. Even after 2.5 months of use, a user reports not knowing how to end a call, and usability matters more over time. When they gain more experience of the product, the role of usability increased presumably as their evaluations of usability became more objective and positive emotions lost their dominating role. This finding offers one possible explanation to the phenomenon that objective usability and subjective usability are often only weakly related [13], but as this study shows the correlation is strengthening over time.

One explanation for the conflicting results is that Karapanos' et al. [21] participants were young, technically oriented and mostly men, whereas our participants came from different ages and backgrounds and were not able to learn or handle all the usability problems with the complicated smart phones. In the other study [20], the product was a simple pointing device, which the participants could quickly learn to use. In sum, the role of usability depends on users, their characteristics and the product. Our novel finding is that the significance of memorable usability problems increases over time.

As two users commented, the time was too short to get to know the new mobile phone and the users may have expected to have problems when they were still unfamiliar with the new phone. Over time, users may become irritated by repeated impracticalities over time as also suggested by Kujala et al. [23, 24]. In addition, positive expectations may have an effect in the early evaluation. For example, Raita and Oulasvirta [32] showed with a short usability test experiment that positive expectations boosted post-test usability ratings even when the users failed most of the tasks.

#### **Validity and implications**

The ecological validity of the study is good. Users used the mobile phones in their own environment in real-life settings. In earlier studies the investigated time periods have been relatively short [24] and performed in experimental settings. Thus, this study provides important insight about the role of emotions over time in real life.

However, as the sample is relatively small and we can't compare reliably for example the importance of memories and actual emotional experiences.

The findings have important implications for the research of user experience. First, they provide empirical support for the claims that user experience is actively constructed rather than a mere collection of emotional reactions. Thus, the whole concept of user experience should focus more on its interpretative and memory dependent nature rather than on immediate reactions of users.

Second, the findings suggest that the importance of memories needs to be considered when evaluating user experience. Many researchers and practitioners find it necessary to assess user experience while a user interacts with a product [25] and many user experience methods focus on momentary experiences or single behavioral episodes [40]. Moreover, retrospective methods have been criticized as unreliable [22]. Yet, the current study suggests that users' retrospective evaluations, while somewhat inaccurate, may be more important than online measurements.

As all experience episodes are not equally important, we need to identify those that are best recalled and most critical to customer satisfaction and loyalty. Several evaluation methods focus on memories (iScale [22] and UX Curve [24]). These can be used to identify what experience episodes are most memorable and meaningful for a given product. The negative issues can be then removed and positive issues can be supported by design.

Third, we need to consider memory and emotions in design. The content analysis of peak and low experience episodes suggest that design should support people to achieve their goals and provide positive experiential consequences. In particular, based on the content analysis of low and peak experience episodes, it looks like that the product should support positive self-image, personalization and have positive social consequences. Further, usability problems that have negative social consequences seem to be especially memorable and avoidable.

Ultimately, designing rewarding and memorable episodes means going beyond ease of use and utility. The product should have a positive influence on the user's personal needs and life.

#### ACKNOWLEDGMENTS

We thank Marlene Vogel, Oscar Person, Merja Halme, Lutz Gegner and anonymous reviewers for their help and valuable comments and Jari Lipsanen and Oscar Person for help with statistical analysis.

#### REFERENCES

- <http://www.attrakdiff.de/en/Home/>
- Bargas-Avila, J.A. and Hornbaek, K. Old wine in new bottles or novel challenges: a critical analysis of empirical studies of user experience. In *Proc. CHI 2011*, ACM Press (2011), 2689–2698.
- Chitturi, R., Raghunathan, R., and Mahajan, V. Delight by design: The role of hedonic versus utilitarian benefits. *Journal of Marketing* 72, 3 (2008), 48–63.
- Csikszentmihalyi, M., & Larson, R. Validity and reliability of the experience sampling method. *Journal of Nervous and Mental Disease*, 175 (1987), 526–537.
- Finstad, K. The usability metric for user experience. *Interacting with Computers* 22, 5 (2010), 323–327.
- Forlizzi, J. and Battarbee, K. Understanding experience in interactive systems. In *Proc. DIS 2004*, ACM Press (2004), 261–268.
- Hassenzahl, M. The interplay of beauty, goodness, and usability in interactive products. *Human-Computer Interaction* 19, 4 (2004), 319–349.
- Hassenzahl, M., Diefenbach, S., and Göritz, A. Needs, affect, and interactive products—Facets of user experience. *Interacting with Computers* 22, 5 (2010), 353–362.
- Hassenzahl, M. and Sandweg, M. From mental effort to perceived usability: transforming experiences into summary assessments. In *Proc. CHI 2004 extended abstracts*, ACM Press (2004), 1283–1286.
- Hassenzahl M. and Tractinsky, N. User experience - a research agenda. *Behaviour & Information Technology* 25, 2 (2006), 91–97.
- Hassenzahl, M. and Ullrich, D. To do or not to do: Differences in user experience and retrospective judgments depending on the presence or absence of instrumental goals. *Interacting with Computers* 19, 4 (2007), 429–437.
- Holland A. C. and Kensinger, E. A. Emotion and autobiographical memory. *Physics of life reviews*, 7, 1 (2010) 88–131.
- Hornbæk, K. and Law, E. L-C. Meta-Analysis of Correlations Among Usability Measures. *Proc. CHI 2007*, ACM Press (2007), 617–626.
- Hox, J. Multilevel analysis. Techniques and applications. Lawrence Erlbaum Associates, New Jersey, 2002.
- Hsee, C.K. and Hastie, R. Decision and experience: why don't we choose what makes us happy? *Trends in Cognitive Sciences* 10, 1 (2006), 31–37.
- Isomursu, M. Tähti, M. Väinämö, S. and Kuutti, K. Experimental evaluation of five methods for collecting emotions in field settings with mobile applications. *International Journal of Human-Computer Studies*, 65, 4 (2007), 404–418.
- Jordan, P.W. Human factors for pleasure in product use. *Applied Ergonomics* 29, 1 (1998), 25–33.

18. Kahneman, D., Krueger, A. B., Schkade, D. A., Schwarz, N., Stone, A. A. A Survey Method for Characterizing Daily Life Experience: The Day Reconstruction Method. *Science* 306, 5702 (2004), 1776-1780.
19. Kahneman, D., Krueger, A. B., Schkade, D. A., Schwarz, N., Stone, A. A. *The Day Reconstruction Method (DRM): Instrument documentation*. Retrieved August 26, 2011, from [http://harris.princeton.edu/faculty/krueger/drm\\_documentation\\_july\\_2004.pdf](http://harris.princeton.edu/faculty/krueger/drm_documentation_july_2004.pdf)
20. Karapanos, E., Hassenzahl, M., and Martens, J.B. User experience over time. In *Proc. CHI 2008 extended abstracts*, (2008), 3561–3566.
21. Karapanos, E., Zimmerman, J., Forlizzi, J., and Martens, J. B. User experience over time: an initial framework. In *Proc. CHI 2009*, ACM Press (2009), 729–738.
22. Karapanos, E., Zimmerman, J., Forlizzi, J., and Martens, J. B. Measuring the dynamics of remembered experience over time. *Interacting with Computers* 22, 5 (2010), 328-335.
23. Kujala, S., Roto, V., Väänänen-Vainio-Mattila, K. and Sinnelä, A. Identifying hedonic factors in long-term user experience. In *Proc. DPPI 2011*, 137-144.
24. Kujala, S., Roto, V., Väänänen-Vainio-Mattila, K., Karapanos, E. and Sinnelä, A. UX Curve: A method for evaluating long-term user experience. *Interacting with Computers* 23 (2011), 473-483.
25. Law, E. L. C. Roto, V., Hassenzahl, M., Vermeeren, A. P. O. S., and Kort, J. Understanding, scoping and defining user experience: a survey approach. In *Proc. CHI 2009*, ACM Press (2009), 719–728.
26. Mahlke, S. Understanding users' experience of interaction. In *Proc. ECCE 2005*, 251–254.
27. Mahlke, S., Thüning, M. Studying antecedents of emotional experiences in interactive contexts. In *Proc. CHI 2007*, ACM Press (2007), 915–918.
28. Miron-Shatz, T. Evaluating multiepisode events: Boundary conditions for the peak-end rule. *Emotion* 9, 2 (2009), p. 206-213.
29. Miron-Shatz, T., Stone, A. and Kahneman, D. Memories of yesterday's emotions: Does the valence of experience affect the memory-experience gap? *Emotion* 9, 6 (2009), 885-891.
30. Mitchell, T.R., Thompson, L., Peterson, E., Cronk, R. Temporal adjustments in the evaluation of events: The "rosy view". *Journal of Experimental Social Psychology* 33 (1997), 421-448.
31. Norman, D. A. Memory is more important than actuality. *Interactions*, March + April (2009). 24-26.
32. Raita E. and A. Oulasvirta, A. Too good to be bad: Favorable product expectations boost subjective usability ratings. *Interacting with Computers* 23, (2011), 363-371.
33. Reichheld, F.F. The one number you need to grow. *Harvard Business Review* 81, 12 (2003), 46-54.
34. Robinson, M. D. and Clore, G. L. Episodic and semantic knowledge in emotional self-report: Evidence for two judgment processes. *Journal of personality and social psychology* 83, 1 (2002), 198-215.
35. Sanders, E.B.-N. and Dandavate, U. Design for experiencing: New tools. *Proceedings of the First International Conference on Design and Emotion, Delft*, (1999), 87–92.
36. Schkade, D.A. and Kahneman, D. Does living in California make people happy? A focusing illusion in judgments of life satisfaction. *Psychological Science* 9, 5 (1998), 340-346.
37. Stone, A.A., Shifman, S.S., & DeVries, M.W. Ecological Momentary Assessment. In D. Kahneman, E. Diener, & N. Schwarz (Eds.), *Well-being: The foundations of hedonic psychology* (pp.26-39). New York: Russell Sage Foundation.
38. Stone, A. A., Schwarz, J. E., Broderick J. E., and Shiffman, S. S. Variability of momentary pain predicts recall of weekly pain: A consequence of the peak (or salience) memory heuristic. *Personality and Social Psychology Bulletin* 31, (2005), 1340-1346.
39. Tuch, A. N., Bargas-Avila, J. A., Opwis, K., and Wilhelm, F. H. Visual complexity of websites: Effects on users' experience, physiology, performance, and memory. *International Journal of Human-Computer Studies* 67, 9 (2009), 703–715.
40. Vermeeren, A.P.O.S., Law, E.L.C., Roto, V., Obrist, M., Hoonhout, J., and Väänänen-Vainio-Mattila, K. User experience evaluation methods: current state and development needs. In *Proc. NordiCHI 2010*, ACM Press (2002), 521–530.
41. Wirtz, D., Kruger, J., Napa Scollon, C. and Diener, E. What to do on spring break? *Psychological Science* 14, 5, (2003), 520–524.
42. Wright, P. and McCarthy, J. *Technology as Experience*. MIT Press, Cambridge, MA, USA, 2004.