

A Wild Evaluation of Users' Emotional Engagement

Petra Sundström^{1,2}, Anna Ståhl², Kristina Höök^{1,2}

¹DSV KTH/SU

Forum 100

164 40 Kista, Sweden

{petra, kia}@dsv.su.se

²SICS

Box 1263

164 29 Kista, Sweden

{petra, annas, kia}@sics.se

INTRODUCTION

The research field of affective computing [7] has been a huge inspiration source for this work, however, being more interested in experiences than efficiency we want to give users a more active role in interacting with such systems. While automatic capturing of users' emotional state has been the most prominent approach within affective computing we instead want to place users very much in the center of a communicational experience. Letting users be engage with and experience their emotions both on a cognitive and a physical level we believe we can create richer emotional experiences for the user. Similarly to an engaging real life conversation where body language and ambiguous emotional expressions are very much a part of what is interesting and appealing with such conversation we want to create an interaction model where the system in a way play the part of an emotionally intelligent conversation partner, similarly to Romero and Mateas ideas of an alien presence [8]. We want our systems to some extent recognize users' emotional expressions but also influence and reinforce expressed emotions. We want the user to be engaged in an *affective loop experience* with the system, where she gets more and more involved in exploring and interacting with her own emotional state.

As a research vehicle we have designed, implemented and evaluated a mobile service we call *eMoto* that aims to embody characteristics of an affective loop experience, briefly outlined above.

EMOTO

eMoto is a mobile messaging service for sending and receiving affective messages [2]. The user expresses her emotions by the use of affective gestures and the system responds to those gestures using colors, shapes and



Figure 1: The eMoto prototype.

animations combined as affective expressions in the background of her text message. Neither gestures nor graphical expressions are, however, completed in the design. The expressions are designed to embody the inner experience of emotional expressions and formed from an analysis of emotional body language [1], but they are not taken all the way to iconic expressions. There are no one-to-one mappings between emotions and expressions; instead they together set up a foundation of expressiveness open for personality, context and situation.

The prototype is built in Personal Java and runs on Sony Ericsson's P800 and P900 mobile phone series. The stylus used with the touch screen on these mobile phones has been extended with an accelerometer, a pressure sensor and Bluetooth to capture and wirelessly communicate the gestures to the mobile phone, see figure 1.

EVALUATION

To determine whether someone is emotionally engaged and moreover emotionally involved in an affective loop, is however a very difficult task. Moreover, we have the difficulty of eMoto being a mobile application and should therefore be experienced and evaluated the wild and not in the more staged lab environment where we believe it is hard for us to create a setting where users can become truly emotionally engaged. But also the wild would not be that much wild, if we researchers came along. Instead we have looked for methods that will let users act as researchers and document their experiences but also to some extent help us in analyzing these experiences. However, to document and analyze a situation when being emotionally engaged with something else we believe is not possible. Most likely you report on emotional engagement after it has occurred and not while it is happening. This and the fact that eMoto so much is inspired by emotional body language, a language in which friends understand each other far better than outsiders do, led us to the idea of having friends to users act as spectators. For the evaluation of eMoto we therefore went for a combination of *Cultural probes* first described by Gaver and colleagues [3] and Isomurso and colleagues' *Experience clips* [5], both methods that contain this aspect of letting users themselves act as researchers.

Cultural Probes

Since Gaver and colleagues article on cultural probes [3] there has been many researchers inspired by this approach. Looking at the different kinds of probes that have been used



Figure 2: The left picture shows the probes given to the user while the right picture shows the probes of the spectator.

there are first of all the probes following directly on the initial idea of using probes as an ethnographical tool, a way of giving people means, typically a diary, a disposable camera and postcards, to report on their individual experiences on a certain topic. Inspired by this method there are now also the more technical probes, using for example mobile phones and text messages to prompt users with questions [4]. In his work on intimacy Kaye has looked into the possibility of using probes as a method for evaluation [6]. However, while Kaye uses his application to get his subjects to think more of and reflect more on their relationships we use probes for evaluation to get users to reflect more explicitly on the design and the affective loop experience we are trying to design for.

Experience Clips

Since eMoto builds very much on a physical experiences and since emotional engagement not is easy to capture in text or in separate pictures we decided for this user study to combine the Cultural Probes method with Experience Clips, a method for mobile evaluation first thought of by Isomursu and colleagues [5]. Similar to the Cultural Probes method Experience Clips also builds very much on the idea of incorporating users more in the design process. What Isomursu and colleagues did was to let users work in pairs and document each other. They let users borrow their mobile phone applications and an additional mobile phone with video capabilities and asked them to focus on feelings, emotions and subjective experiences, aspects that are very hard to capture with traditional methods. They also let users take turn in being the user or the spectator. However, due to storage capabilities of the film strips users only got to borrow the mobile phones for a maximum of two hours. In our user study of eMoto we very much want to come past novel usage and strive for results on users' appropriation and more personal experiences and therefore a longer evaluation time is needed. In our user study users got to use eMoto and document their usage for a time period of two weeks.

Procedure

The evaluation took place over a period of two weeks, in September 2005.

We gave users probe bags containing a notebook, some postcards, a disposable camera and the eMoto prototype

itself consisting of a Sony Ericsson P910 mobile phone with an extended stylus, but also we prepared bags with a notebook, some postcards and a small video camera for the users to give to a close friend of theirs who would act their spectator, see figure 2. Both bags also contained letters with indications for what kind of things we were expecting of them as users or spectators and suggestions for how to go about their different tasks of exploring from the inside vs. exploring from the outside.

After the two weeks we quickly looked through all the films, notebooks and postcards from both users and their spectators and together with our log data formed a quick and dirty analysis of each user's experiences. We then discussed this analysis with our users during individual interviews. The intention was to get a starting point for the discussion but still stay to the notion of letting the users act as researchers. During the interviews users were also asked to go through and explain/interpret all emotos they had sent and received, they divided them into two groups; *authentic emotos* that were said to express an emotional value and *tests* that were sent either to explore the possibilities of the service or only because the users took part in this evaluation.

Subjects

From a previous lab study of eMoto [10] we had got to know Agnes¹ who seemed to be the perfect user for a system of this kind. According to our user-centered design process where the focus on a specific user group is key and due to the fact that we previously had looked at more usability aspects of eMoto [9,10], the fact that Agnes had previous knowledge of the system was only positive at this stage. Since eMoto very much builds on friends and their additional knowledge of each other we wanted Agnes to use the system together with a few of her friends. We had five prototypes so Agnes could ask four of her friends to participate, the five young women that took part in this evaluation where then Agnes, Isabella, Louise, Susie and Mona:

Agnes, a 25 years old Behavioral scientist living in Uppsala (the 4th largest city in Sweden, approximately 100 km

¹ All names of the users are fictional.

outside Stockholm) had her boyfriend as her spectator and it was only she who actually knew all the other young women in the user group.

Isabella, a 26 years old Consultant in the energy sector is an old friend of Agnes and she did not know any of the others. Isabella lives a busy working life in Stockholm and since she does not live with anyone she asked if she could use two spectators, her cousin and one of her friends, which we agreed on.

Louise knew and still knows only Agnes and Susie, however, Susie is one of her better friends and living alone she asked Susie's roommate and spectator to be also her spectator. At the time of the evaluation she was 25 years old and wrote her Master thesis in Technical Biology at the Royal Institute of Technology (KTH) in Stockholm.

Susie and her roommate, however, stay in Uppsala and it was at the University there that Susie similarly to Louise wrote her Master thesis in Technical Biology. She is 24 years old and as said above she had her roommate as her spectator. Before the user study she knew Agnes and Louise, Louise a bit better though, but during the two weeks she not only spent more time than she had done before together with Agnes but she also got to know Mona quite well.

Mona, 25 years old, is the third user from Uppsala. She is an Economics school graduate but at the time of the evaluation she as Agnes still took a few courses and worked extra at one of Sweden's phone companies. Living together with a few of her friends also she had her roommate as her spectator.

RESULTS

The results of this evaluation are of three kinds; results on eMoto as a mobile service, results from an affective loop perspective and results on methodology. In here we mainly focus on the method of having users acting as researchers as a method for evaluating emotional engagement and involvement in an affective loop, but first a short summary of the overall picture.

Table 1 shows how much of the probes that we returned from each user and also what else there was for each person in the log data. In total 96 emotos were sent during the two weeks; 78 of them were sent between the users themselves and 18 they sent to friends who were not part of the user group. After the five users had interpreted the 78 emotos sent between them we could complete the communicational puzzle for these emotos. The emotos sent to friends outside the user group has not been interpreted only explained by the user who sent them. What we found was 24 *successful emotional messages*, authentic emotos that also were understood by the receiver, 4 *misunderstandings*, authentic emotos which were interpreted having a different emotional value than what was intended by the sender, 24 as a consequence of *novel vs. more mature usage*, emotos that were either sent or interpreted as authentic emotos but

placed in the test pile by either the sender or the receiver, and finally 24 emotos that were placed in the test pile by both sender as receiver.

	Log data and returned probes
Agnes	33 sent emotos and 28 received, 38 SMS left in memory, 5 postcards, ~8 minutes of experience clips, her and her spectator's notebooks
Isabella	16 sent emotos and 10 received, 0 SMS (she unfortunately had to carry also her regular phone), 2 postcards, ~7 minutes of experience clips, her and her spectator's notebooks
Louise	11 sent emotos and 10 received, 2 SMS left in memory, no postcards, no experience clips and no notebooks
Susie	13 sent emotos and 13 received, 13 SMS left in memory, 2 postcards, ~10 minutes of experience clips, her notebook but not her spectator's
Mona	23 sent emotos and 17 received, 11 SMS left in memory, 3 postcards, ~11 minutes of experience clips, her notebook but not her spectator's

Table 1: Each user's contribution.

An Affective Loop Experience

First of all, our idea of building a foundation of expressiveness based on characteristics of emotional body language seemed to have been a successful design approach, Mona says:

"I leave out things I think are implicit due to color... the advantage is that you don't have to write as much, it's like a body language. Like when you meet someone you don't say 'I'm sulky' or something like that, because that shows, I don't need to say that, and it's the same here, but here it's color."

Users were able to express themselves and their personality in the expressions but also they understand each other's personal language, and this even though some of them did not know each other so well:

Mona: *"All these colors are kind of perky and cheerful, I would guess Isabella, if I have green and Agnes has yellow, then perhaps she has orange to express that she's a little bit happy and perky."*

Figure 3 shows Isabella, Mona and Susie engaged with the eMoto prototype. Not are these pictures from experience clips signs of emotional engagement but also we see these pictures as examples of personality as well in the gestures, this even though we can not compare the separate gestures with each other since we do not know what the three users exactly are engaged in expressing at these specific movements.

Apart from the experience clips there are also comments on emotional engagement in the notebooks, for example Agnes spectator writes:

"When she was happy she showed that with her whole body, not just her arm was shaking but her whole body while a huge smile appeared on her lips."



Figure 3: Isabella, Mona and Susie engaged with the eMoto prototype.

Having Users Acting as Researchers

Furthermore, having users acting as researchers was indeed a successful approach. All except for Louise seemed to be very positive in getting the responsibility, Isabella said:

“Great fun, interesting... especially to let users be so much in focus with the prototype themselves.”

To get users more emotionally engaged we had in previous user studies of eMoto noted that letting them act out in a scenario was very beneficial for how much they were able to relax and enjoy themselves [10]. Writing scenarios that are true for each user is however not an easy task. As a consequence of letting users themselves set the scenarios in this evaluation of eMoto we noted that what is true to the users is not always what we think it is. One evening during the two weeks the three users living in Uppsala, Agnes, Susie and Mona met for what they called an *eMoto-pub*, where they met in a bar, discussed the prototype and sent emotos to each other and to the other two users living in Stockholm.

Furthermore, as researchers we often draw the wrong conclusions when we look at data from user studies. As the two weeks past by it became more and more apparent to us that Louise did not feel comfortable with the eMoto service. Looking at especially one emoto she sent to Agnes therefore surprised us. This was an emoto were she had truly expressed her love for Agnes, something we did not think was something Louise would do, but after having discussed this with both of them it became known to us that Agnes had become quite annoyed with Louise incapability of not sharing her emotions with others and especially with her who was a friend so Agnes therefore stole Louise’s phone and sent this warmhearted emoto to herself as if it was from Louise. To us it was not only the fact that Louise did not fit our targeted user very well but also we saw that this method of letting the users act as researchers was very uncomfortable to her. However, to us this fact was more interesting than disturbing, not only do we wish to design our prototypes for a targeted user group but now also we know that we need to keep our users in mind when exploring and choosing evaluation methods.

Since users were able to use eMoto in their daily lives for as long as two weeks we think we received data on a much deeper level than for example comments on usability and

first impressions. Four of our five users, however, still thought the two weeks were too short. All except for Louise explicitly stated they wanted to keep eMoto after the two weeks were over.

Having a spectator for two weeks was on the other hand a bit difficult. The spectator could of course not be there all the time and was perhaps not there when the most exciting experiences occurred, something that Susie was worried about, but still she liked the idea of having a spectator:

“It was fun also. Because first I thought that if one shall take videos, then perhaps one... in other words, one always behave affectedly on video, but when one sat there with the phone and the stylus it became like... one forgot about the camera, in any case I did that. Then she asked questions and I thought it became... it felt serious.”

From the spectators point of view it became a problem that they were not users themselves. For a similar user study in the future we would have to work more on the incentives of spectators. Perhaps we should take inspiration from Isomursu and colleagues’ work where they took turns in being user and spectator. On the other hand, we still want to keep the notion of getting an outsiders point of view especially since privacy and questions like ‘what shall others say when I do this in public?’ is one of the very interesting aspects of physical interaction in a mobile and public environment.

Furthermore, what users see as researchers needs to be what we want feedback on; if not we have to educate our users before they can help in performing the task of analyzing themselves. For this user study users saw eMoto as a mobile service for sending and receiving affective messages, they did not see a prototype exemplifying an affective loop experience and could therefore not comment on such an experience. From previous user studies of eMoto in the laboratory we have rewarding results on the interaction model [10], out in the wild, on the contrary, the gestures were more regarded as an obstacle in finding the right graphical expressions. Between the lines, though, in their notebooks and in their experience clips there are signs of the users being emotional engaged with the gestures, however, if we want users to be able to act as researchers and see these signs themselves we would need to design more explicitly for what we want users to see and give us feedback on, or educate them on our research questions.

If we bring our prototypes out in their intended environment they will be exposed to greater expectations than in an evaluation taking place in a more staged environment, such as the laboratory. The further we take our prototypes towards being evaluated as products the further we have to bring them to a level of actually being products. In the eMoto case we had worked hard and iteratively with users in the laboratory on refining the affective expressions [9,10] but aspects of our prototype that do not answer directly to our research questions we had not worked as much with. Our research approach is to take our prototypes to a level where they exemplify our research questions. Closer to real usage this level seems to be much higher than in the laboratory. In our previous user studies on eMoto users seemed to have been able to look past these details to a much higher level, out in the wild, however, their acceptance level was much lower. All five users in this wild evaluation of eMoto had problems with the form factor of the extended stylus for example:

Mona: *"The stick is too big & tall & looks like a dildo. I know I shall try to look beyond the technical, but it undeniably restrains usage."*

SUMMARY

We do think we have got results here indicating that we are truly heading in the right direction in finding a high-quality evaluation model for emotional engagement. Having users act as researchers out in their own environment and not in the more staged laboratory we strongly believe is a beneficial approach to take when looking for something as subjective and personal as emotional engagement.

REFERENCES

1. Fagerberg, P., Ståhl, A. and Höök, K. Designing gestures for affective input: an analysis of shape, effort and valence, In Proceedings of Mobile Ubiquitous and Multimedia, MUM 2003, Norrköping, Sweden, 2003.
2. Fagerberg, P., Ståhl, A. and Höök, K. eMoto – Emotionally Engaging Interaction, Journal of Personal and Ubiquitous Computing, Special Issue on Tangible Interfaces in Perspective, Springer, 2004.
3. Gaver, B., Dunne, T., and Pacenti, E. Cultural probes, In Interactions, Volume 6, Issue 1, pp. 21-29, ACM Press, New York, USA, 1999.
4. Hulkko, S., Mattelmäki, T., Virtanen, K., and Keinonen, T. Mobile Probes, In Proceedings of NordiCHI'04, Tampere, Finland, 2004.
5. Isomursu, M., Kuutti, K., and Väinämö, S. Experience Clip: Method for User Participation and Evaluation of Mobile Concepts, In Proceedings of Participatory design 2004, Toronto, Canada, 2004.
6. Kaye, J. The logbook probe: facilitating auto-thick description for evaluation, Submission for Workshop on Innovative Approaches for Evaluating Affective Systems, Stockholm, Sweden, 2006.
7. Picard, R. Affective Computing, MIT Press, Cambridge, MA, USA, 1997.
8. Romero, M. and Mateas, M. A Preliminary Investigation of Alien Presence. In the Proceedings of Human-Computer Interaction International (HCII 2005), Las Vegas, NV, USA, July 2005.
9. Ståhl, A., Sundström, P. and Höök, K. A Foundation for Emotional Expressivity, DUX 2005, San Francisco, CA, USA, 2005.
10. Sundström, P., Ståhl, A. and Höök, K. eMoto - Affectively Involving both Body and Mind, In Extended abstract CHI'05, Portland, Oregon, USA, 2005.