EmoSnaps: A Mobile Application for Emotion Recall from Facial Expressions

Evangelos Niforatos
Madeira Interactive Technologies Institute
Campus da Penteada.
Funchal, 900-390, Portugal
evangelos.niforatos@m-iti.org

Evangelos Karapanos
Madeira Interactive Technologies Institute
Campus da Penteada.
Funchal, 900-390, Portugal
ekarapanos@uma.pt

CHI'13, April 27 – May 2, 2013, Paris, France.
ACM 978-1-XXXX-XXXX-X/XX/XX.

Abstract
We introduce EmoSnaps, a mobile application that captures pictures of one's facial expressions throughout the day and uses them for later recall of her momentary emotions. We describe a field study that employs EmoSnaps in an attempt to investigate if and how individuals infer emotions from self-face pictures. Recalled emotions through EmoSnaps are compared to ground truth data as derived from Experience Sampling. Contrary to our expectations, it was shown that by increasing the temporal distance of recall to an experience and by reducing temporal context, we increase users' ability to infer their emotions from a self-face picture.

Author Keywords
Experience sampling, Day reconstruction method, diary studies, affective computing.

ACM Classification Keywords
H.5.2 Information interfaces and presentation (e.g., HCI): User Interfaces: Evaluation/Methodology

Introduction
The increasing emphasis on how mobile technologies are experienced in everyday life has resulted in an increased interest in in-situ measurement and, in particular, the Experience Sampling Method (ESM) [12]. ESM is considered as the gold standard of in-situ

measurement [8] as it samples experiences and behaviors right at the moment of their occurrence, thus reducing memory and social biases in self-reporting. However, ESM entails important drawbacks, such as disrupting users' activity and imposing high burden to them [16].

Motivated by these drawbacks, Daniel Kahneman and colleagues proposed the Day Reconstruction Method (DRM) [8], a retrospective self-report protocol that aims at increasing users' accuracy in reconstructing their experiences at the end of a studied day. It does so by imposing a chronological order in reconstruction, thus providing a temporal context for the recall of each experience. DRM has been found to provide a reasonably good approximation to experience sampling data [8] and the method has been well adopted also in the HCI community (see [10] for a review). In our line of research we attempt to contribute towards a next step in the field of momentary assessment, that of *technology-assisted reconstruction* (TAR) [10]. TAR consists of passively logging users' behaviors throughout the day with mobile sensor technology and employing these data to assist the reconstruction of one's daily activities and experiences.

This work introduces EmoSnaps, a mobile application that captures unobtrusively pictures of one's facial expressions throughout the day and uses them for the later recall of her momentary emotions. We present a two-week-long deployment of EmoSnaps with five participants that inquired into *if* and *how* self-face pictures assist the reconstruction of momentary emotions.

Emotion and Facial Expressions

Emotions are so tightly connected with facial expressions that one could even question whether there can be emotion without facial expression [5]. Not only it is difficult for people to hide their emotions in facial expressions, research has also shown that humans are surprisingly accurate in recognizing basic emotions, such as anger, disgust, fear, joy, sadness and surprise, from facial expressions [13,15]. Especially when it comes to happiness, research has revealed that humans can accurately recognize the emotion in 96.4% and 89.2% of the times in Western and non-Western cultures, respectively [16].

Algorithmic techniques in emotion recognition have flourished [1,3,14] and provide a promising approach in stationary settings. On the contrary, mobile settings introduce substantial complications in capturing facial expressions. Some novel solutions have been proposed by Teeters, Kaliouby and Picard [17] on "Self-Cam", a chest-mounted camera that is able to detect 24 feature points on the face and extract emotions using dynamic Bayesian Models, as well as Gruebler and Suzuki [6] on a wearable interface device that can detect facial bioelectrical signals. While providing the ability to capture emotions in a continuous fashion, both these approaches are highly intrusive, inducing a feeling of being monitored as well as raise concerns of social acceptance, especially when long-term deployments in real-life settings are concerned. With EmoSnaps, we aimed at creating a tool that can be truly transparent in daily life and can be employed in long-term field studies.

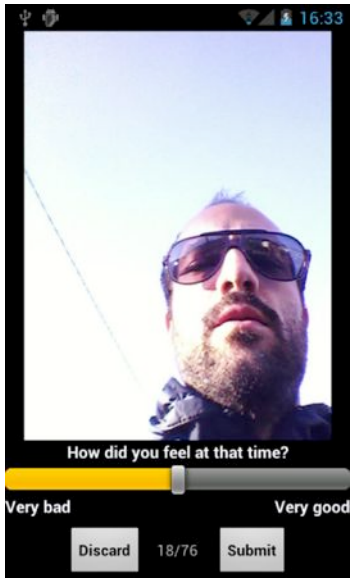

Figure 1. EmoSnaps captures unobtrusively self-face pictures throughout the day. It then provides these pictures to the user in an attempt to assist him or her in reconstructing his or her momentary emotions.

EmoSnaps

EmoSnaps employs an event driven sampling, where predefined events, such as "screen unlock", "phone call answer" and "sms sent" trigger a photo snap. Upon triggering, a picture using the front facing camera of the mobile device is captured within 300-500 milliseconds, thus being truly transparent to the user. A custom-made interface (Fig. 1) is used in revisiting all captured pictures and reconstructing one's momentary emotions. EmoSnaps runs on Android OS and has been tested and optimized for efficiency so far on the Nexus S mobile device.

Study

We conducted a two-week-long deployment of EmoSnaps with seven participants to inquire into *if* and *how* self-face pictures assist the reconstruction of momentary emotions. First, given prior literature [13,15], one would expect participants to be able to accurately recognize their own emotions given a self-face picture. However, a less accurate emotion recognition could be expected due to the mobile setting, as pictures may be of varying orientation, luminosity and image quality.

Second, if participants can accurately recognize their emotions in self-face pictures, the question is how do they do so? We can think of at least two ways (Fig. 2). The first assumes that individuals will recognize their emotions through their facial expressions in the self-face picture [2,15]. In contrast, the second assumes

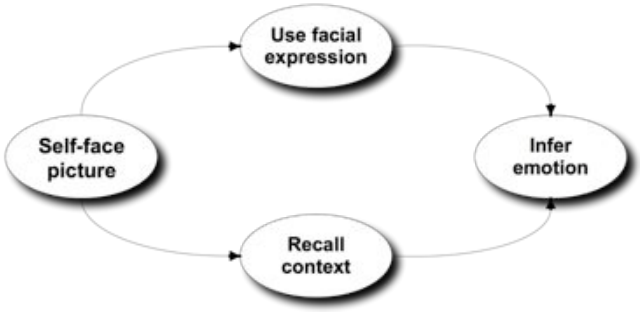

Figure 2. If people accurately recognize their emotions, how do they do so? Inferring emotions directly from facial expressions, or recalling episodic memories and drawing upon this knowledge to infer emotion?

that individuals will use cues of the picture to recall episodic memories (e.g., where they were, what they were doing, who they were with) and, based on this information, to recall their emotions at that given time. Recent work has suggested that *"emotional experience can neither be stored nor retrieved"*, but can only be reconstructed on the basis of recalled contextual cues from episodic memory [9]. If self-face pictures are recent and contain information that may cue episodic memories, participants could as well infer their emotions from these episodic memories rather than infer them from facial expressions of the picture. One could expect these *reconstructed emotions* to be more accurate than the *recognized emotions*, given that participants may draw upon rich episodic information in the case of a recent event.

Third, one might suggest that the relevant others of an individual are better able to infer the individual's emotions from her facial expressions given increased

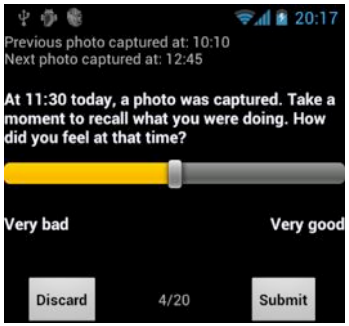

Figure 3. Time-based reconstruction was employed as a control condition.

exposure to them. Indeed, research has shown that when participants are subjected to a task of identity matching, reaction times to familiar faces are faster than reaction to unfamiliar faces. Yet, there is no difference in reaction time between familiar and unfamiliar faces in tasks of facial expression matching [18]. Others have, however, shown that familiarity may make a difference by improving the accuracy in recognizing emotions [11]. Given these results, we expect familiar others to be better able to infer the individual's emotions from her facial expressions.

Study Design

To address these three research questions, we formed four conditions, each representing a distinct reconstruction process as follows.

PHOTO-DAY RECONSTRUCTION

At the end of each day participants are asked to revisit all self-face pictures taken throughout the day and recall how they were feeling at the time of each captured picture (Fig. 1). Pictures are presented in chronological sequence as this has been proven to enhance the reconstruction of episodic cues [1]. Thus, we assume participants in this condition to have access to both approaches of emotion inference, *recognition* and *reconstruction*.

TIME-DAY RECONSTRUCTION

At the end of each day, participants are asked to recall what they were doing at the time when a picture was captured and recall how they were feeling at that time (Fig. 3). The times of the preceding and the succeeding pictures are shown as it might provide a temporal context and, thus, assist the reconstruction process [8]. As in the Photo-Day condition, picture times are presented in chronological sequence but not the actual pictures. This type of reconstruction serves as a control condition and any difference between this and Photo-Day reconstruction in terms of participants' accuracy will be attributed to the effect of the self-picture.

PHOTO-WEEK RECONSTRUCTION

A week after the last day of the study, participants are asked to revisit all pictures taken in the Time-Day condition and recall how they were feeling at the time when each picture was captured. They use the same interface as in the Photo-Day condition (Fig. 1), but this condition differs in two respects. First, as a week or more has elapsed since these pictures were taken, we assume participants to be unable to reconstruct episodic memories related to the picture. Second, pictures are presented in random order in an effort to minimize any effect of building contextual knowledge as participants go through the pictures. Thus, in this condition, we assume participants to infer their emotions only from facial expressions.

PHOTO-RELEVANT RECONSTRUCTION

For each participant, a relevant other is chosen to evaluate the same pictures the participant has evaluated during the Photo-Week reconstruction. Relevant others consisted either of the partners-in-life, or, the closest colleague of each participant. We judged that both these groups of individuals would have an increased familiarity with participants' facial expressions. Relevant others used the same interface as in Photo-Day and Photo-Week reconstructions (Fig. 1) with the pictures being displayed in random order.

Figure 4. ESM was used throughout the day to self-report on momentary psychological wellbeing.

Figure 5. Clustering Areas of Interest (AOIs) for eye tracking analysis for each picture.

Measures

Motivated by previous work in the field of self-report on psychological well-being [4], we designed a simple interface (Fig. 4) to inquire into participants' happiness at c1/10/13 12:46 AMertain moments. By employing ESM, we asked participants to quantify their happiness using a continuous scale, ranging from 1 (Very bad) to 100 (Very good), using a widely validated single-item construct of psychological wellbeing [4]. The same bar was also used during all the reconstruction sessions. The difference Δ between the self-reported emotion during Experience Sampling and during reconstruction signifies participants' inaccuracy in reconstruction.

Eye-tracking was used to understand participants' eye gaze behaviors in the different reconstruction conditions. Each picture was preprocessed so that two major Areas of Interest (AOI) are defined, Face and Background (Fig. 5). Due to high effort in defining AOI, we selected a random sample of 10 rated pictures per condition (Photo-Day, Photo-Week and Photo-Relevant) for each participant. Two metrics, *visit count* (i.e., number of eye fixations to an AOI following a fixation to a different AOI) and *total visit duration* (i.e., total time in seconds a participant spent looking at a specified AOI) were used to indicate eye gaze behaviors.

Behavioral logs were analyzed to understand how participants interacted with the interface when asked to infer their emotion. We analyzed 4 metrics in total: (1) *Photo duration*, the total time taken to evaluate a picture in seconds, (2) *Number of scale interactions*, the number of times the user interacted with the rating scale for each picture rating, and (3) *Duration of scale interactions*, the total duration (in milliseconds) of all interactions with the rating scale for each picture rating (the duration of each interaction was estimated from the difference of the drag and drop events of the scale). Finally, *total delta* reflects the summative variance of the scale cursor movement during these interactions.

Retrospective Think Aloud (RTA) [7]sessions were finally conducted to obtain qualitative insights in the way participants and their relevant others infer emotions from facial expressions. RTAs were performed for all three conditions that include facial expressions as cues (Photo-Day, Photo-Week and Photo-Relevant). For this purpose, an RTA protocol was formed mainly questioning the rationale behind emotion inference.

Participants

Seven individuals (2 females, median age 29 years) participated in the study for a total of six days. All were office workers with similar work patterns. They all used the application during working days.

Procedure

Each participant was given a Nexus S mobile device with the EmoSnaps application installed. Each time a participant unlocked the screen a picture would be taken and the participant would be prompted to self-report on psychological wellbeing using a validated single-item continuous scale [4] (Fig. 3). This would run for a total of six days (three days in Photo-Day and three in Time-Day, order counterbalanced across participants). One week after, participants would perform the Photo-Week reconstruction and their relevant others would perform the Photo-Relevant reconstruction.

Figure 6. Heat map produced by summarizing gaze behavior on a self-face picture.

Figure 7. Interface used for performing Retrospective Think Aloud (RTA) sessions.

All reconstruction sessions were performed on the Tobii Eye Tracker T60, running an Android OS emulator at the size of Nexus S mobile device. The relevant others also repeated the Photo-Relevant reconstruction session on the eye tracker with the same configurations. During the evaluation on the eye tracker, a video was capturing both the participant's eye-gaze behaviors as projected on the screen as well as the participant's face (see Fig. 7). Upon completion, both participants and their relevant others underwent a Retrospective Think Aloud (RTA) [7] session, cued by the captured video.

Results

In this section we present the results categorized according to our research hypotheses. The previously described measures are combined in order to explain and complement the phenomena observed.

Emotion inference from self-face pictures

A total of 584 pictures were captured in the course of the six days. Participants and relevant others were able to infer emotions for approximately 70.6% of the pictures. For the remaining 29.4%, they clicked on the "Discard" option. As participants reported, this happened primarily due to poor lighting conditions, privacy concerns, incorrect posture, or inability to infer one's emotions from their facial expressions.

"[P1] I discarded it because it was blurry and poor. I wouldn't do it if the photo was looking silly, but I would do it for privacy reasons", "[P2] I am not really expressive in these pictures.", "[P4] It's always the same! Looks like I don't have a happy face! It's a family problem I guess!"

Unexpectedly, participants sometimes reported that the aesthetics of their self-face pictures influenced the way they inferred emotions, while in some cases they reporting discarding a picture because they did not like their appearance:

"[P4] This one looks nice! The photo looks nice so I was feeling happy!", "[P5] That's the thing of being a girl again, I look at the picture and I am like oh I have such a huge nose! So am not sure it's kind of a girl thing but it is inevitable for me to not look at these kind of things sorry!"

Discard rates ranged per condition with the highest discard rate observed in Photo-Week (36%), followed by Photo-Day (35.4%), Photo-Relevant (34.7%) and Time-Day (2.3%). A Pearson chi-square analysis between the Photo-Week and Time-Day conditions on discard rates revealed a significant difference between the two distributions ($X^2(1, N = 659) = 99.83$, $p < .001$). Thus, the sample is reduced to 1002 valid pairs of emotion ratings (ones coming from experience sampling and ones from reconstruction sessions). On average, participants would capture a total of 15 pictures in a given day (min = 8, max = 29).

An analysis of variance on use logs, with photo duration, number of scale interactions, duration of scale interactions and total delta as dependent variables and type of reconstruction (Photo-Day, Time-Day, Photo-Week, Photo-Relevant) as independent variable, displayed a significant main effect for the type of reconstruction on the total delta ($F(3,419) = 14.298$, $p < .01$, $h^2_p = 0.092$). Post-hoc tests using the Bonferroni correction revealed that participants in the Time-Day ($M = 18.798$, $SD = 13.403$) condition displayed a significant higher total delta than in Photo-Day ($M = 13.292$, $SD = 7.417$) and Photo-Week ($M =$

13.048, *SD* = 8.019) conditions ($p < .05$). This may indicate that the absence of self-face picturea as a cue implies greater uncertainty when trying to infer emotion, thus taking more time to decide on one's rating.

Context Recall vs. Facial Expression Use

An analysis of variance with the z-transformed computed distance *Δ* between experience sampling and reconstruction values as dependent variable and type of reconstruction (Photo-Day, Time-Day, Photo-Week, Photo-Relevant) as independent variable, displayed a significant main effect for the type of reconstruction ($F(3,998) = 4.553$, $p < .01$, $h^2_p = .014$). Post-hoc tests using the Bonferroni correction revealed that participants in the Photo-Week (*M* = 9.722, *SD* = 9.629) condition were significantly more consistent in estimating their emotion as compared to the Photo-Day (*M* = 12.242, *SD* = 11.857) condition ($p < .05$).

An analysis of variance with visit count and total visit duration as dependent variables and type of reconstruction (Photo-Day, Photo-Week, Photo-Relevant) and AOI (Face, Background) as independent variables, displayed a significant main effect for the type of reconstruction and for the AOI on visit count ($F(2,72) = 4.251$, $p < .05$, $h^2_p = .106$). Post-hoc tests using the Bonferroni correction revealed that participants in Photo-Day (*M* = 1.240, *SD* = 2.067) condition had significant higher visit count on the Background AOI than they had in the Photo-Week (*M* = 0.320, *SD* = 0.627) condition ($p < .05$). In agreement with our expectations, this indicates that in Photo-Day condition participants relied more on the context of the picture to infer their emotion, as compared to the photo-Week condition where context cues might have been lost due in memory.

Indeed, the majority of the participants in Photo-Day condition repeatedly reported emotional inference primarily based on their location, their social context and the activity they were engaged into, while partially neglecting facial expressions.

> "*[P2] I know I was feeling pretty well because I was eating… You know that feeling when you are close to the tree and you eat more fruits than you actually eat at home.*", *"[P5] … I know I was having lunch, because I know this is next to the bar. So I know I was feeling good because we were with Leonardo talking and making jokes so I know I was OK."*

Often, these inferences were based on patterns of behavior and experience rather, thus potentially introducing recall biases [20]: *"[P4] I can tell that because it is always a good time having breakfast with my colleagues all together"*

One participant explained how she used contextual cues to infer her emotions through her pictures: *"[P3] I don't relate the context to emotion directly. I look at the context to recall what I was doing and by what I was doing I can recall if I was happy."*

However, the context remained the same during the day, it was reported of secondary importance: *"[P5] I am pretty sure I took all the pictures at home so maybe the background is kind of secondary to me so I know where I was all the time."*

Participants also reported that the presenting pictures in their temporal order supported the process of inferring their emotion: *"[P2] The sequence of the photos helps as I can understand what I was doing."*

In contrast to Photo-Day condition, the Photo-Week revealed an opposite effect. All participants reported emotional inference based on their facial expressions captured in self-face pictures. Facial expressions were preferred over context in multiple cases.

"[P3] This one I cannot tell, am still at work from the context, but I don't see the mouth and I cannot really tell by the eyes so I discarded it.", "[P4] I only concentrate on my face… this shows if you are happy or not… and maybe the time but the face comes first."

Participants also described the areas of their face on which they concentrated the most during reconstruction. One's mouth, eyes and eyebrows were the most frequently reported.

Relevant other versus Self

Post-hoc tests using the Bonferroni correction revealed that relevant others in Photo-Relevant (M = 12.091540, SD = 9.590864) condition were significantly less consistent as compared to participants in Photo-Week (M = 9.722, SD = 9.629) condition ($p < .05$). Interestingly however, participants in Photo-Day condition displayed a significant greater *total delta* than relevant others did in Photo-Relevant (M = 11.333, SD = 6.813) condition ($p < .05$). This indicates that relevant others were significantly more certain when evaluating an individual's face picture than the individual was at the end of the day. This could be explained by the fact that relevant others ignored the context of the pictures they were evaluating. In fact, Photo-Day condition had a significant higher visit count on the Background AOI than the relevant others in Photo Relevant (M = 0.320, SD = 0.556) condition ($p < .05$). This effect is complemented by relevant others reporting that they relied totally on individual's facial expressions to infer emotion from pictures of familiar faces.

"[R4] I can't understand what she was doing by the pictures in none of them."

In cases where neither a distinctive facial expression, nor any particular contextual cue was present, relevant others reported difficulty in their rating:*"[R3] I have no indication, I have no memory, he didn't come up to me to tell me how he was feeling so I try to guess, that way it makes it a lot harder."*

Interestingly, relevant others used cues of individual's behaviors to infer their emotional state: *"[R3] So basically he is with his sunglasses on. He is pounding with his lips this is something he does when he is not in a very good mood."*

Discussion

Overall, the results are rather surprising and contradict our a-priori expectations, in that the Photo-Week produced a significant higher accuracy in emotion inference when compared with Photo-Day. In other words, when reviewing the self-pictures a week after they were captured, participants could more accurately infer their emotions than when reviewing them at the end of the day. Eye tracking analysis revealed that participants in Photo-Day condition displayed a significant greater number of visit counts on the

background part of the picture when compared with Photo-Week and Photo-Relevant conditions. In the Photo-Day condition participants focused more times on the background than in the Photo-Week. A possible explanation is that at the end of the day participants tried to recall contextual information in order to infer their emotions while neglecting their facial expressions. This observation is also supported with participants reporting that in Photo-Day condition contextual information derived from the Background was used to infer activity and subsequently emotion.

Surprisingly, individuals also proved to be significantly more accurate than relevant others in inferring their emotions. Both in Photo-Week and Photo-Relevant condition, a reluctance of context utilization can be assumed for different reasons. On the one hand, participants in Photo-Week experience a one-week long interval between the experience and its reconstruction. On the other hand, context is meaningless for the relevant others as showed in the RTA results. Although both participants and their relevant others used the same areas of the face (Mouth, Eyes and Eyebrows) to infer emotion in Photo-Week and Photo-Relevant, respectively, the relevant others proved to be less accurate. A possible methodological explanation for this is that even though z-transformation was applied to normalize the Δ between ESM and Reconstruction ratings, relevant others did not have a notion on the spectrum used by the participants.

Interestingly, the usage log analysis revealed a significant higher total delta for Photo-Day condition when compared with Photo-Relevant. This possibly reveals a higher degree of uncertainty for participants inferring emotion at the end of the day than their relevant others, likely to the use of a higher number of cues. These findings provide support for the recognition and not the reconstruction approach. One possible explanation for this phenomenon could be that the process of inferring emotions from reconstructed episodic memories conflicts with the one of inferring them from facial expressions, thus disrupting the recognition process. An alternative possible explanation could be a learning effect in the Photo-Week, as participants were more familiar with the reconstruction interface since they used it before in the Photo-Day condition.

Conclusion

In this paper we proposed EmoSnaps, a mobile application that captures unobtrusively pictures of one's facial expressions throughout the day and uses them for later recall of her momentary emotions. In a two-week long field deployment of EmoSnaps we attempted to inquire into *if* and *how* self-face pictures assist the reconstruction of momentary emotions. Contrary to our a-priori expectations, we found that emotion recall with self-face pictures takes place through inferring them from facial expressions and not through a reconstruction of episodic memories. Specifically, it was shown that by increasing the temporal distance of recall to an experience and by reducing temporal context, we increase users' ability to infer their emotions from a self-face picture.

The significance of these findings needs to be noted as it suggests that designers, contrary to common sense, should avoid employing EmoSnaps or related emotion recognition approaches based on self-face pictures for recent experiences, but rather employ this to "recall" experiences that lie further in the past. We consider

this a considerable strength of the approach as users may "recall" their emotions for long periods of time, in one-go.

References

[1] Anderson, S.J. and Conway, M.A. Investigating the structure of autobiographical memories. *Journal of Experimental Psychology: Learning, Memory, and Cognition 19*, 5 (1993), 1178.

[2] Azcarate, A., Hageloh, F., Van de Sande, K., and Valenti, R. Automatic facial emotion recognition. *Universiteit van Amsterdam*, (2005).

[3] Cohen, I., Sebe, N., Garg, A., Chen, L.S., and Huang, T.S. Facial expression recognition from video sequences: temporal and static modeling. *Computer Vision and Image Understanding 91*, 1 (2003), 160–187.

[4] Conner, T.S. and Reid, K.A. Effects of intensive mobile happiness reporting in daily life. *Social Psychological and Personality Science 3*, 3 (2012), 315–323.

[5] Ekman, P. Facial expression and emotion. *American Psychologist 48*, 4 (1993), 384.

[6] Gruebler, A. and Suzuki, K. Measurement of distal EMG signals using a wearable device for reading facial expressions. *2010 Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, (2010), 4594 –4597.

[7] Guan, Z., Lee, S., Cuddihy, E., and Ramey, J. The validity of the stimulated retrospective think-aloud method as measured by eye tracking. *Proceedings of the SIGCHI conference on Human Factors in computing systems*, (2006), 1253–1262.

[8] Kahneman, D., Krueger, A.B., Schkade, D.A., Schwarz, N., and Stone, A.A. A Survey Method for Characterizing Daily Life Experience: The Day Reconstruction Method. *Science 306*, 5702 (2004), 1776–1780.

[9] Karapanos, E., Martens, J.B., and Hassenzahl, M. Reconstructing experiences through sketching. *arXiv preprint arXiv:0912.5343*, (2009).

[10] Karapanos, E. Beyond Experience Sampling: Evaluating Personal Informatics with Technology-Assisted Reconstruction. *arXiv preprint arXiv:1207.1821*, (2012).

[11] Kenny, D.A. and Acitelli, L.K. Accuracy and bias in the perception of the partner in a close relationship. *Journal of Personality and Social Psychology; Journal of Personality and Social Psychology 80*, 3 (2001), 439.

[12] Khan, V. J. and Markopoulos, P. *Experience sampling: A workbook about the method and the tools that support it*. Eindhoven, 2009.

[13] Ortony, A. and Turner, T.J. What's basic about basic emotions? *Psychological review 97*, 3 (1990), 315.

[14] Pantic, M. and Rothkrantz, L.J.M. Automatic analysis of facial expressions: The state of the art. *Pattern Analysis and Machine Intelligence, IEEE Transactions on 22*, 12 (2000), 1424–1445.

[15] Russell, J.A. Is there universal recognition of emotion from facial expressions? A review of the cross-cultural studies. *Psychological bulletin 115*, 1 (1994), 102.

[16] Scollon, C.N., Kim-Prieto, C., and Diener, E. Experience sampling: Promises and pitfalls, strengths and weaknesses. *Journal of Happiness Studies 4*, 1 (2003), 5–34.

[17] Teeters, A., El Kaliouby, R., and Picard, R. Self-Cam: feedback from what would be your social partner. *ACM SIGGRAPH 2006 Research posters*, (2006), 138.

[18] Young, A.W., McWeeny, K.H., Hay, D.C., and Ellis, A.W. Matching familiar and unfamiliar faces on identity and expression. *Psychological research 48*, 2 (1986), 63–68.

[20] Schwarz, N., Kahneman, D., Xu, J., Belli, R., Stafford, F., & Alwin, D. (2009). Global and episodic reports of hedonic experience. Using calendar and diary methods in life events research, 157-174.