# Experience in Social Affective Applications: Methodologies and Case Study

**Paul André**

ECS, University of Southampton

United Kingdom, SO17 1BJ

pa2@ecs.soton.ac.uk


**m.c. schraefel**

ECS, University of Southampton

United Kingdom, SO17 1BJ

mc@ecs.soton.ac.uk

**Alan Dix**

Computing Department

Lancaster University, UK

alan@hcibook.com


**Ryen W. White**

Microsoft Research

Redmond, WA, USA

ryenw@microsoft.com

## Abstract

New forms of social affective applications are emerging, bringing with them challenges in design and evaluation. We report on one such application, conveying well-being for both personal and group benefit, and consider why existing methodologies may not be suitable, before explaining and analyzing our proposed approach. We discuss our experience of using and writing about the methodology, in order to invite discussion about its suitability in particular, as well as the more general need for methodologies to examine experience and affect in social, connected situations. As these fields continue to interact, we hope that these discussions serve to aid in studying and learning from these types of application.

## Keywords

Methodology, design, evaluation, experience, affect, social, social networking, well-being.

## ACM Classification Keywords

I.m Computing Methodologies: Miscellaneous.

## General Terms

Design, Measurement, Theory

## Introduction

There is an emerging interest at CHI in social networks and connection [9,22], as well as in experience of use and affect [11,12,15]. We're interested in how these two strands might be combined more formally to investigate how the group sharing of personal state, expressed in terms of physical and emotional well-being, might enhance quality of life in terms of social interaction and engagement, even creativity and innovation. Since Twitter, Facebook, etc. have opened up the desire to chat, to announce, to convey the personal to public or social groups, we are investigating whether well-being representation can be formalized and used constructively to support the aforementioned quality of life metrics. There are two key aspects to this work: the concept and design of a tool to share well-being information, and evaluating the experience of how that information is perceived, interpreted and used.

While we have designed an artefact to explore these exchanges, some key questions emerge. What is an appropriate evaluative methodology when trying to make sense of a non-task based experience that, unlike previous applications, focuses on unambiguous group social settings? What happens when you are both interested in the experience *and* the artefact that facilitates that experience, i.e., both design evaluation and evaluation of experience?

This paper is an exploration of how we assessed the variety of methods available to us, devised a hybrid evaluation approach, and met resistance from reviewers on questions of validity. The goal of this paper is to generate discussion towards finding an appropriate evaluative solution for what may be a potential new kind of application: when designing for experience and affect meets with personal and social networking for a particular goal, in this case social support for better engagement and quality of life.

The rest of the paper offers an application overview, situates it in the context of related work, elaborates on our methodology, arguing for and against each aspect, and closes with opportunities for future work.

## Healthii: Self-Reflection and Group Awareness of Well-being Online

Motivated by psychology research suggesting benefits in assessing subjective well-being both individually [7], and within a group [5], and a lack of rich or explicit well-being in online practice, we developed Healthii, a social networking tool for conveying well-being. The tool lets us explore whether enabling the expression of well-being status within social networking sites would be perceived as valuable, and how that utility was
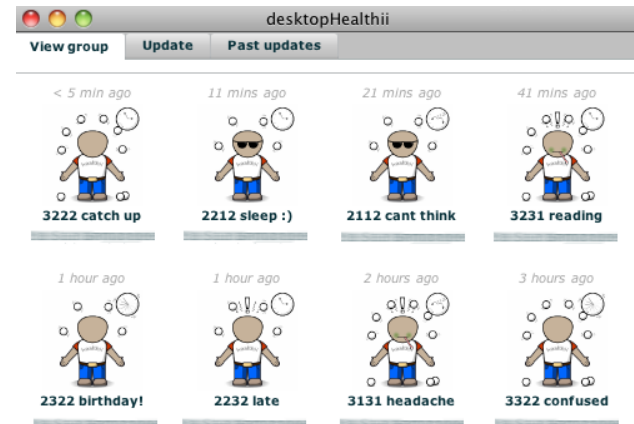


Figure 1. An example of a group view of participants, showing both numeric and avatar view.

perceived: in the ability to express well-being, in self-reflection, in group awareness, or all of the above. In the interests of space we skip a complete design rationale (see [1]), and highlight the features of the system.

Healthii uses a set of four discrete dimensions (busy, enjoyment, stress and health), and three finite values (not, quite, very) within those dimensions to reflect personal wellbeing. A person's wellbeing status can be represented by an avatar or a numeric code. Users can update their own state and view their past states, or view their friend group, as seen in Fig 1. Updating Healthii can be achieved through Facebook (or a desktop application) by radio button (see Fig 2), or through Twitter. Updating via Twitter involved adding the hashtag #healthii, and then encoding one's state into the numerical representation, for instance #healthii(3222:CHI!) would represent 3=very busy, the

three 2s for "usual" enjoyment, stress and health, and "CHI!" as the reason.

## Related Work

Healthii is focused on the experience of reflecting and sharing physical and emotional well-being. As such it draws from two fields in particular: user experience and affective computing that:

- try to address human needs beyond the instrumental, stressing affective and emotional aspects of the interaction [11],

- move away from a task-based analytical approach to design systems to support rich, meaningful and pleasurable experiences [18],

- and/or support people in understanding, interpreting and experiencing emotion in its full complexity and ambiguity [3,12].

Two such systems in particular are relevant, both in their function and in evaluation. Affector [19] is a video window between the neighboring offices of two friends to communicate their moods, the images distorted based on sensor readings (e.g., movement in the office) and user mappings. Similarly, the eMoto [21] system is designed for expression of affect in mobile phone text messages, allowing users to alter the background color and pattern of their message with gestures, conveying how the sender is feeling through pressures, movement patterns or pace. With Affector, the researchers used themselves as designers, users and evaluators of the system, and in eMoto five friends used the emotional messaging service for two weeks, acting as both users and spectators that "observe and document user behavior". Both studies found that more than just conveying a simple emotion at a specific time, the open-ended expressions allowed creative use and emotional meanings to emerge over the course of interaction, with the relationships outside of the system putting meaning to and affecting the implications within the system. In the following section, we explain how Healthii differs from these approaches.

*On clarity and ambiguity*

Our goal has been to see how concision and constraints on expressing complex internal states via a specific vocabulary of terms such as "bored, sick, really busy, feeling great" can be used functionally. Such discrete dimensions are distinct from recent work in emotional computing, which has a similar goal to Healthii of understanding, reflection, and awareness of a variety of mood and emotion. While work in this area has encouraged flexible interpretation of mood and emotion [19,21], such ambiguity of expression is mostly used in a rich 1-to-1 context, where choice of a certain word or colour carries personal connotation. Where there has been an appeal to a wider group, more interpretive



Figure 2. Input of four well-being dimensions is via radio button, and represented by an avatar and numeric status.

methods tend to focus on encouraging the reflection of the individual, and in some cases ambiguity in public/group scenarios has led to a misunderstanding of the original meaning [3]. It is less clear how to harness heuristics such as ambiguity in the case of trying to allow some assessment of "group mood" unless there were to be some emergent group conventions.

The pre-coded answers that we use trade off individual expressive flexibility for ease of group comprehension, maintaining a level of global consistency and transparency. Constrained discrete dimensions also meant we could take advantage of embedding an encoded textual status into social networks. By using these simple discrete scales early on, we can reduce the drain on "emotional effort" incurred when being thoughtful about representing oneself [6], and perhaps move to more complex representations as people develop suitable self-expression skills [17].

## Method
In tracing the history of evaluation, Kaye & Sengers [15] conclude that evaluation is not simply a methodological issue, but an epistemological one, i.e., in arguing what knowledge is and how it is generated. Similar to situating HCI into three (epistemological) paradigms [10], they state that recently there has been a push towards the: *"rejection of cognitive approaches based on the modeling, reducibility or predictability of human behavior. [Instead seeing] the world as locally produced by its members in an ongoing fashion, which cannot be adequately represented by a formal model."*

This latest paradigm contains a variety of perspectives and approaches whose central metaphor is: all action, interaction, and knowledge is seen as embodied in situated human actors [8,10]. In other words, the way we come to understand the world, ourselves, and interaction derives from our location in a physical and social world, highlighting subjective experience. This perspective clearly influences methodologies for design and evaluation of technologies that address issues in this paradigm, in particular, we are investigating how to model that experience and enhance our understanding of the world, towards improving quality of life with applications such as Healthii.

Before we can consider enhancing quality of life however, we must understand:

1)  Whether the approach we propose in Healthii is used and valued (i.e., the expression and viewing of self and group status via constrained dimensions)?

2)  In what way is that value or utility experienced?

As such, we were interested in not exactly the tool itself (it is one of many possible instantiations), but of evaluating the concept of wellbeing expression, seeking to understand the experience of use to determine what value there is.

As we have seen, the evaluation of experience is *not* novel. It falls into Ramage's third type of CSCW evaluation, the "conceptual development" to evaluate the concepts that underlie a system, and recent work has focused on the experience of use, user feedback, and discussions that surround a tool [3,15,20,21]. Our approach is also similar to Technology Probes [13]: simple, flexible, adaptive technologies with three goals: the social science goal of understanding the needs and

desires of users in real-world settings, the engineering goal of testing the technology, and the design goal of inspiring users and researchers to think about new technologies.

What *is* novel is both the concept being studied (as explained, affective computing to date has tended towards interpretive flexibility rather than the constrained dimensions we use towards pragmatic effect), and our specific methodology for experience in social affective applications. We summarize our methodology in the next section, highlight two critiques, list the aspects we considered to be vital in exploring our questions, before discussing existing methodologies and presenting arguments for and against specific points of ours.

*Chosen Methodology*
We selected ten people from our lab who were friends or colleagues and users of social networking tools. This helped to ensure that we were not creating a new friend group and could concentrate on the effect of Healthii. Participants were given the Healthii tool (our artefact), and asked to use it over the course of five weeks. During those five weeks, we would meet as a group once a week to discuss how people were using the tool, share experiences or anecdotes, and positive or negative aspects. To facilitate an optimal experience for the participants, we were open to reconsidering certain aspects of the design of the tool or its interaction with Facebook/Twitter throughout the trial. An online survey was conducted at the end of the trial to gather individual feedback.

We have struggled to define an 'acceptable' method for evaluating such applications. We think the hybrid we present is effective, but are concerned about the anticipated (or already experienced) critique, and so offer an argument/counter-argument approach to explain and examine our methodology. We aim to promote discussion around the question: for current and potential future experience-focused social applications, unique considerations seem to call for a new methodology. We think we considered the issues and chose an appropriate evaluation, but... were we wrong?

Though we borrow from a number of existing methodologies (longitudinal studies, diary studies, and participatory design for instance), our methodology is also distinct in several respects motivated by the phenomena we were exploring. Based on these phenomena and questions, and relevant to social or experience-focused applications in general, we considered the following aspects of an evaluation to be vital:

- Awareness of how usage affected a participant personally (usually achieved through diary studies, or interviews)

- Awareness of how usage affected perception or action towards other participants, i.e., any group affect (usually achieved through group interview, or to some extent, participatory design)

- Tailoring of both the concept and tool to fit the user's desired experience (or removal of barriers to that experience) (depending on the stage of the design, usually achieved through various probes, participatory design, design based research)

- Promotion of discussion (either with the evaluator, or between participants) as to the experience of use, and the goals participants had.

- A sufficient period of evaluation to achieve the above sort of experience to allow in-depth discussion (a longitudinal study). Also relevant is the ability to live with the tool, to enable that level of engagement.

## Critiques of the Methodology

In the reviews for our CHI paper [1], two critiques in particular were made. In essence they were that combining the design iterations of participatory design with a longitudinal study cast doubt on the validity of results, and that due to the mix of methods it was unclear whether the paper was about design or about evaluation. We emphasize that we are not claiming our paper was unassailable, but rather that because of the interesting and contentious discussion around these different methods, we hope to engage the community in a discussion of appropriate methods for evaluating potentially new types of application in experience-focused social settings. We discuss the two critiques in more detail here, before comparing our methodology to existing ones, and presenting arguments for and against specific parts of our methodology.

*Refining design.* Reviewers disliked that, as previously mentioned, we were open to refining the design to optimize experience throughout the trial, seeing it as an uncomfortable mix of participatory design and longitudinal study. Aside from any fault in the way we presented the method which may have resulted in ambiguity or misinterpretation, was it an inappropriate choice? We argue that we were exploring a concept,

and were not tied to a particular instantiation where changing design may affect a measured outcome. Thus, when we received feedback that the tool was confusing or hindering the desired experience, it seemed reasonable to change the tool to fix these problems. The changes did not affect the concept of discrete well-being we were interested in, and not making the changes would have meant a frustrated participant group for the remainder of the study.

*Design phase.* Since we were familiar with our target group and had related work to draw upon, we based our design on interviews and the literature, deciding against other methods (such as cultural probes or true participatory design) because we had a clear idea of a concept we wished to investigate. These methods, along with a previous study [2], gave us both the dimensions of well-being (busy, engagement, stress, health), and the visual design. A critique of the reviewers was that this was too cursory, and more of a focus on the elements of design would have been appropriate. Though these dimensions were not rigorously tested or iterated on, they were borne from interviews and related work, and discussed regularly and in-depth during the deployment. One such dimension (engagement) after weeks of use was deemed to interact too much with 'busy', and so was replaced with 'enjoyment', which enabled a different aspect to be conveyed. Such changes did not affect the concept of discrete well-being, but were able to improve the experience of what the participants wanted to convey, and were only brought to light through our use of a longitudinal study.

## Arguments for the Methodology

In this section we first discuss why existing methodologies didn't seem to fit the goals of our project, before an explicit outlining of the rationale for (and against) each contentious part.

*Comparison to Existing Methodologies*

- *Participatory design* didn't exactly fit our goals, as we had a particular design concept in mind (testing constrained well-being input), and an artefact we had developed after conducting interviews and related work searches. Participatory design, on the other hand, can involve stakeholders earlier on in the design process, in the analysis of needs for instance, and evolve the design and artefact [16]. We discuss participatory design and probes in Future Work. However, we were interested in the optimal experience for participants, and so were open to changes and refinements to the artefact which enhanced the experience without altering the concept or questions we were interested in.

- Proponents of *longitudinal studies* would take issue with the above idea of altering the artefact in the midst of the trial. We present arguments for and against this in a later section.

- *Ethnography*. By purely observing and monitoring usage, and potentially interviewing one-on-one, we do not think we would have fully explored the awareness of effect on the group, nor promoted the discussion we hoped for. Similarly to above, observing an artefact in use also traditionally precludes altering that artefact.

- *Focus groups* promote the idea of value in seeing people discuss an idea or artefact, and so we were keen to utilize this approach.

- *Methodologies for experience or affect.* Taking inspiration from the evaluation of tools like eMoto, we considered how interpreting data along with the participants was necessary and allowed insights into experience.

These considerations left us with a list of desires for evaluation that while taken individually were not new, no methodology exactly catered to them all. These were close contact with our participants, close contact between participants to discuss the data and experience of use, and the ability to refine the artefact over time. In some ways the overall question seemed to be: could we take some of participatory design and focus group techniques and apply them to a longitudinal study? Below we describe our methodology, and focus on each potentially contentious aspect.

*How the Methodology Aided Our Findings (or Didn't)*

In this section we discuss each aspect of the methodology, presenting arguments for and against, and considering how each aspect helped answer our questions about the affect of a social tool. We started with an artefact we had developed in a previous trial that was robust to use, as a sufficient starting place to get at the discussion about experience.

WEEKLY GROUP MEETINGS

Argument Against: Weekly Group Meetings

By meeting as a group, some people may dominate or lead the discussion, or not speak due to feeling

uncomfortable sharing something. Demand characteristics may also pose a problem, participants unconsciously changing their behavior or assuming a role to help or hinder the experiment.

Argument For: Weekly Group Meetings
Rather than use the update type emails/questionnaires that might be distributed to individuals during a longitudinal field trial, we instead drew on participatory design models to have our participant group meet and discuss their experiences with the tool. We thought this hybrid approach reasonable because of our particular interest less in the artefact itself and more on the experience enabled by it. Because the tool was inherently social, and participants saw and interacted with all the available data in usage of the tool, and even saw each other socially, we felt it was not a problem to discuss the tool as a group. Indeed, we considered it a benefit. We feel that the group meetings allowed an open and in-depth discussion of experience, with anecdotes or suggestions sparking other people's imaginations or memories. Because of the exploratory nature of the study, participants were not privy to our hypotheses, and so demand characteristics in that regard were not an issue. It is feasible that they ascribed value when none was apparent, or followed others' leads, but participants did not seem shy to contradict one another, or discuss a dislike of aspects of the tool. We argue that in these sorts of affective computing scenarios, as Technology Probes states, these artefacts reject the strategy of collecting 'unbiased' ethnographic data, but we reap the benefits of collecting data in-situ. We discuss the use of individual data collection below.

## DESIGN REFINEMENTS
Argument Against: Design Refinements
Changing the tool changes the experience, especially halfway through a longitudinal field study.

Argument For: Design Refinements
Although we did have a novel design of discrete well-being representation, it was the concept or idea we were testing, not making claims about a particular embodiment. We were instead focused on the effect on users, so that we may understand if there was interest and value in the concept. Thus, we recognize that changing the artefact affected the experience, but by only changing the design based on feedback and consensus, it allowed us to explore what the users really wanted of such a system, and how they used it.

We made four updates or alterations to the design over the five weeks. Two were simply related to re-tweeting when someone made an update via Facebook. The other two were deeper and concerned the dimensions and the meaning of numbers. We were pleased that as participants became familiar with the tool, they switched from learning and thinking about how to express a state, to really considering what they were stating and how it was perceived, uncovering some of the deeper experience we wished to explore. These discussions resulted in a change to the meaning of the numbers of the dimensions, our participants reporting such changes made both updating and viewing statuses easier to comprehend.

## LONGITUDINAL FIELD STUDY
We used a longitudinal (or rather, long-term) field study because Healthii introduced new behaviors - recording explicit well-being attributes - and we wanted

to provide sufficient time to see how the application would be used beyond an initial training period. A longitudinal field study, we hoped, would give us this perspective. The value of running the study for five weeks became particularly apparent when by week three participants moved from talk of how each were using the tool to the meanings being conveyed in the dimensions. We had wanted to probe this level of experience rather than the artefact. The longer study time allowed the tool to become transparent enough to focus on that experience.

INDIVIDUAL SURVEYS

To mitigate any discomfort sharing in a group environment, as well as to gain a final individual understanding of experience, we deployed an individual (online) survey at the end of the study to help articulate these reflections.

*Results*

Though the focus of this paper precludes a full discussion of our results, we were encouraged to find participants reporting anecdotes and instances of value in self-reflection at the time of update as well as over time, in group awareness, and a desire to continue using the tool after the trial. Our original paper can be examined for full results [1], though we hope to significantly rewrite based on feedback from reviewers, and hopefully discussion from this paper.

## Discussion Points and Conclusions

Applications focused on affect and experience of use are finding new ground in social, connected settings (see a workshop at this CHI [22], for instance). We are exploring how the rise of social networking can be evaluated through affective and experience metrics,

and how such applications might be tuned to enhance quality of life for better social engagement, creativity, even innovation. Both broad and specific questions are open for discussion. For instance, the validity of knowledge gained through such subjective types of evaluation, the metrics for associating better social awareness with engagement or creativity, and as we increasingly communicate online, how to investigate meaning and affect, both for the particulars of our study as discussed here (e.g., small co-located groups, systems for long-term use) as well as in other situations. There is also a meta-level of discussion surrounding the way such evaluations are received by the community and reviewers.

Taking a well-being application as a case study, it was our aim in this early study to understand if and how discrete well-being would be used by participants. We described how we felt existing methodologies did not enable us to explore the questions particular to these new types of social affective applications. We constructed a hybrid methodology combining group discussion, design refinement and longitudinal field study. We hope that we have explained our reasoning for (and potential arguments against) our approach that, with inevitable time and resource constraints, offered insight into the experience of use of Healthii in both personal and social situations. We see clear directions from our results of what to look at next, and perhaps only the next study, perhaps more traditionally deployed over a social networking site, will vindicate the approach of this one. The current study had ten participants co-located (and two elsewhere). In the future if we, or others, wished to explore affect on disparate groups of people, we may take inspiration from Kaye's Virtual Intimate Object [14], which used

logbooks with open-ended questions to explore context and experience of use.

We emphasize we are not proposing this particular methodology as a panacea for all social or affective applications, but rather to inform a discussion on whether and what new types of methodologies are needed, and what attitudes to them exist, in order to study and learn from them as these fields continue to interact.

## References

[1]   André, P., schraefel, m. c., Dix, A. and White, R. W. Healthii: Self-Reflection and Social Awareness of Wellbeing Online. Technical Report, ECS, University of Southampton. http://eprints.ecs.soton.ac.uk/18350/

[2]   André, P., schraefel, m. c., Dix, A. I feel so 1323 today: Capturing, Portraying, and Interpreting Well-being Online. Poster paper at *UIST'09*.

[3]   Boehner, K., DePaula, R., Dourish, P., and Sengers, P. Affect: from information to interaction. In *CC '05*.

[4]   Boehner, K., DePaula, R., Dourish, P., and Sengers, P. 2007. How emotion is made and measured. *Int. J. Hum.-Comput. Stud.*, 65(4):275–291.

[5]   Brewer, M. and Gardner, W. 1996. Who is this "We"? Levels of collective identity and self representations. *Journal of Personality and Social Psychology*, 71(1):83–93.

[6]   Brown, T. J. 2009. Self informatics. In *CHI'09 Workshop*: Designing for Reflection on Experience.

[7]   Diener, E., Sapyta, J. J., and Suh, E. 1998. Subjective well-being is essential to well-being. *Psychological Inquiry*, 9(1):33 – 37.

[8]   Dourish, P. 2004. *Where the action is: the foundations of embodied interaction*. The MIT Press.

[9]   Grace, J., Zhao, D., boyd, d. Microblogging: What and How Can We Learn From It? *CHI 2010 Workshop*.

[10] Harrison, S., Tatar, D., and Sengers, P. The three paradigms of HCI. In a*lt.chi 2007*.

[11] Hassenzahl, M. and Tractinsky, N. 2006. User experience: a research agenda. *Behaviour and Information Technology*, 25(2).

[12] Höök, K. 2008. Knowing, communication and experiencing through body and emotion. *IEEE Trans. Learn. Technol.*, 1(4):248–259.

[13] Hutchinson, H., et al. Technology probes: inspiring design for and with families. In *CHI 2003*.

[14] Kaye, J. 'J'. I just clicked to say I love you: rich evaluations of minimal communication. a*lt.chi 2006*.

[15] Kaye, J. and Sengers, P. The Evolution of Evaluation. In *alt.chi 2007*.

[16] Kensing, F. and Blomberg, J. 1998. Participatory Design: Issues and Concerns. *CSCW* 7, 3-4.

[17] Lane, R. D. *Cognitive Neuroscience of Emotion*, 2002 Oxford University Press.

[18] Sengers, P. 2004. *The engineering of experience*, in Funology pp. 19–29. Kluwer, Norwell, MA, USA.

[19] Sengers, P., Boehner, K., Warner, S., and Jenkins, T. 2005. Evaluating Affector: Co-Interpreting What 'Works'. In *CHI 2005 Workshop* on Innovative Approaches to Evaluating Affective Interfaces.

[20] Strong, R. and Gaver, B. Feather, scent and shaker: Supporting simple intimacy. *CSCW*. 1996.

[21] Sundström, P., Ståhl, A., and Höök, K. 2007. In situ informants exploring an emotional mobile messaging system in their everyday practice. *Int. J. Hum.-Comput. Stud.*, 65(4):388–403.

[22] Visser, T., van Bel, D., Dadlani, P., Yarosh, S. Designing and Evaluating Affective Aspects of Sociable Media to Support Social Connectedness. *CHI 2010 Workshop*.