Using Visualizations to Increase Compliance in Experience Sampling

Gary Hsieh, Ian Li, Anind Dey, Jodi Forlizzi, Scott E. Hudson
Human Computer Interaction Institute
Carnegie Mellon University, Pittsburgh PA 15213
{garyh, ianl, anind.dey, forlizzi, scott.hudson}@cs.cmu.edu

ABSTRACT
Experience sampling method (or ESM) is a common data collection method to understand user behavior and to evaluate ubiquitous computing technologies. However, ESM studies often demand too much time and commitment from participants, which leads to attrition and low compliance among participants. We introduce a new approach called *experience sampling with feedback* or *ES+feedback* that improves compliance by giving feedback to participants through various visualizations. Providing feedback to users makes the information personally relevant and increases the value of the study to participants, which increases their compliance. Our exploratory study shows that *ES+feedback* increases the compliance rate by 23%.

Author Keywords
Experience sampling, Visualizations, Empirical studies

ACM Classification Keywords
H5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous.

INTRODUCTION
Experience sampling method or ESM (also, ecological momentary assessment [1]) is a common data collection method. ESM supports evaluations that are *in situ* (in the actual situation), ecologically valid (approximates the real-life situation of what is being studied), and span multiple days, thus it is well suited for evaluating ubiquitous computing applications, which are used in different kinds of environments and contexts. In the past few years, ESM has been utilized by many research projects, both to understand user behavior in the field [1,9] and as a means to evaluate ubiquitous computing systems [2,3,6].

However, running a study using ESM is not without problems. One of the biggest problems with ESM studies is that it often requires a lot of time from participants [9].

When a participant is signaled to self-report, even if scheduled, participants have to stop their current task to respond. Furthermore, sampling is triggered multiple times a day for several days, even weeks, creating boredom and irritation on the part of participants. Such demands can lead to low compliance whereby participants ignore self-report prompts.

One solution to the compliance problem is compensation. Using a variable payment scheme based on compliance, participants may be more motivated to respond to sampling requests. Another solution is to use mobile technology to facilitate ESM data collection. Instead of paper and pencil, PDAs [5] and mobile phones [8] have been used to record data. Context-aware devices can trigger sampling at more opportune times [6] and can reduce time spent on self-reports by automatically recording certain data [4].

We propose a way to improve compliance in ESM studies using feedback mechanisms. We call this approach *experience sampling with feedback* or *ES+feedback*. For example, if the study collects self-reports about availability, an *ES+feedback* study could provide visualizations of data for the participant to peruse and explore. We believe that showing participants their own collected information has two key benefits: it makes the information personally relevant and interesting [7] and it increases the value of the study to the participants. Together, these benefits increase participant compliance.

To test the *ES+feedback* approach, we conducted a 25-day field study. In a setup that mimics a typical ESM study, we found that users who saw visualizations of their own data were more likely to respond to sampling requests compared to users who did not see visualizations. The compliance rate of those receiving feedback was *23% higher* than the rate of those who did not receive feedback.

While future work is needed to explore the details of this approach, these findings show that feedback can improve how ESM is currently used. Not only would conducting *ES+feedback* studies help researchers and practitioners collect more data from their participants, participants may also gain from participating in this kind of studies by allowing them to observe something about themselves.

ESM IN UBICOMP AND VISUALIZATIONS

Several ubiquitous computing systems have been developed to facilitate ESM studies. The context-aware experience sampling method [6] reduces demand for input from participants by only collecting self-reports when a particular context is detected. The MyExperience system extends context-triggered ESM with passive logging of contextual information discerned from device usage and sensor readings [4]. The system collects more information without increasing participation from participants. Momento [2] leverages text messaging (SMS) and media messaging (MMS) supported by most mobile phones to collect ESM samples and to share data between participants and experimenters.

Visualizations as feedback have been shown to promote behavior changes [10]. Visualizations engage users to explore and learn about their environment, their social network, and themselves [11]. We believe that we can leverage the ability of visualizations to engage users to increase compliance in ESM studies.

STUDY OF *ES+FEEDBACK*

We conducted a study to explore the feasibility of the *ES+feedback* approach. The study tests whether providing visualizations of the data collected during an ESM study would increase compliance (percentage of samplings completed). We expect that users who receive feedback would be more compliant than those who do not receive feedback. Additionally, when asked to fill out experience sampling questionnaires about information that they do and do not see visualizations for, we expect that participants would respond more to questionnaires that have visualizations than those that do not. Our hypotheses are:

H1 Participants who received feedback would have a higher rate of compliance.

H2 Participants would respond more to the types of questionnaires for which they have feedback.

Method

We designed a deceptive ESM study so that participants are unaware of our goal to see how feedback affects their compliance rate. On our recruitment email and fliers, we told participants that the goal of the study was to model their application usage behavior and to explore the correlation between their application use and mood and between their application use and interruptibility. We called the study "Predicting Interruptibility and Mood by Application Use".

On the first day of the study, participants came to our lab with their laptops. We explained that we were conducting an exploratory study intended to find out if any correlations exist between computer application use and interruptibility, or between computer application use and mood. They were told that during the study, they would receive two types of questionnaires, either about interruptibility or about mood.

Both also ask about their computer application use. These questionnaires would appear at random intervals. The study lasted for 25 days and participants were paid $25 at the end of the study. They were also told that while they are not required to respond to all the questionnaires, it would be better for the study if they completed as many as possible.

Participants were randomly assigned into three conditions, the *Control* condition and two feedback conditions: *A+I* and *A+M*. The *A+I* group saw visualizations on interruptibility, while the *A+M* group saw visualizations on mood. We had two feedback conditions to understand the impact of different types of information (interruptibility, mood) on response rates.

All participants read the consent forms and had the sampling software installed on the their laptops. For the *Control* condition, nothing more was said. For the other two feedback conditions, participants were told that someone within our research department is doing work on visualizations and has completed a display on either application use and mood or application use and interruptibility (depending on their condition), and that we thought it would be nice to allow the participants to see the information collected. They were given instructions on how to access this visualization page and we stepped through an example with them.

After 25 days, participants returned to our lab and completed a questionnaire about their experience with the sampling software. In addition to this questionnaire, participants in the feedback groups completed a questionnaire about the visualizations. Participants were compensated for their time and debriefed about the actual purpose of the study. We also removed the software that was installed at the start of the study.

System

The *ES+feedback* system has two primary components: (1) a sampling software that prompts participants for self-reports at random times and (2) a web site that stores and
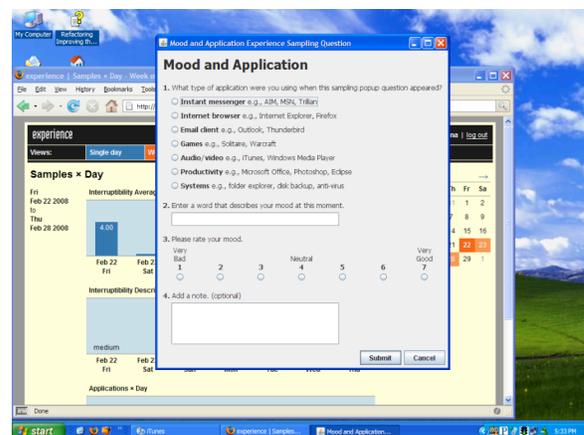

Figure 1. Example of a *Mood* questionnaire appearing while the user is using the computer.

Figure 2. Screenshot of one of the visualizations (By Mood).

visualizes the responses from participants.

The sampling software triggers a questionnaire on the user's computer screen about 10 times a day at random intervals (Figure 1). If the participant chooses not to answer or was not in front of the computer to answer, the questionnaire disappears after a minute. The software stores data about what applications are currently being used on the participant's computer. We use this data during analysis to determine if the questionnaire was not answered because the participant was not actively using the computer or because the participant chose not to answer.

All three conditions saw two kinds of questionnaires: a *Mood* questionnaire or an *Interruptibility* questionnaire. Both questionnaires were equally likely to appear.

The *Mood* questionnaire asks users 4 questions: (1) the type of application they were using when the questionnaire appeared, (2) a word/phrase that describes their mood, (3) a mood rating between 1 (very bad) to 7 (very good), and (4) a note about what they were doing when the questionnaire appeared (optional).

The *Interruptibility* questionnaire is similar to the *Mood* questionnaire, but questions 2 and 3 are rephrased to ask about interruptibility: (2) a word/phrase that describes their interruptibility, *e.g.*, stressed, deadline, or available, and (3) an interruptibility rating between 1 (busy) to 7 (available).

Answers to the questionnaires are sent to a web site for storage. After submitting answers, a dialog box appears to the feedback participants (*A+I* and *A+M*), giving them an option to view their current and past collected information. Feedback participants can also access their collected information at any time by clicking on the application icon
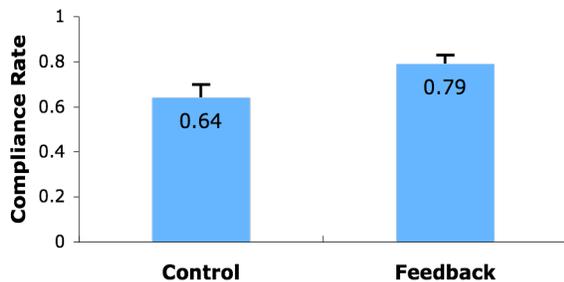
in the system tray. In the *Control* condition, a dialog box does not appear after participants submit their responses. The *Control* group did not know about the web site and did not use it.

On the web site (Figure 2), the feedback participants can view several visualizations of their mood or interruptibility by day, week, period of the day (morning, afternoon, evening, midnight), application usage, and average ratings.

Participants

We recruited 26 participants using a recruiting web site and newsgroups (*A+M* had 10, *A+I* and *Control* each had 8). All participants owned a laptop computer that ran Windows XP and had Internet access at home or at work.

Results and Discussion

Compliance rate is defined as the number of completed responses divided by the total number of times the questionnaires appeared while the user was using the computer. The total number of questionnaires includes the times when participants chose not to answer the questionnaires either by ignoring them or by clicking the "Cancel" button on the questionnaire window, but does not include the times when the user was not present at their computer. The analysis looked at the data by manipulation (*Control*, *A+I*, *A+M*) and by whether the group received feedback (*Control* and *Feedback*).

Hypothesis 1: *Feedback leads to higher compliance rate*
For this analysis, we combined *A+I* and *A+M* into a single *Feedback* group and compared it to the *Control* group, the difference is significant (79% to 64% respectively, $F(24)=4.71$, $p=0.04$) (

Figure 3).

To compare compliance trends, we analyzed compliance rate by 5-day periods. We expect that as the study progressed, participants would respond to fewer questionnaires. Our analysis shows a significant interaction effect on compliance rate between the groups and periods [$F(2,2)=4.11$, $p=0.02$]. In Figure 4, compliance rate for the *Control* group drops off as the study progressed while the compliance rates of the other groups remain about the same.
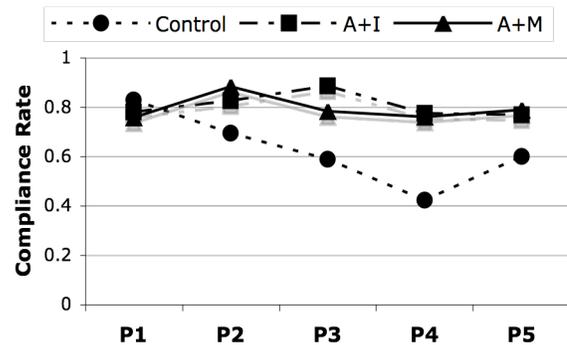
In sum, we have strong support for Hypothesis 1, *i.e.*,

Figure 3. Compliance rates of *Control* and *Feedback* groups.

Figure 4. Compliance rates of each group by period.

receiving feedback increases compliance rates. The increase is 23%. Additionally, the *Feedback* group better sustained their compliance rate throughout the study, while the *Control* group decreased their rate.

Hypothesis 2: *Feedback groups will respond more to questionnaires for which they have visualizations*

We expected that the value of providing users feedback regarding their personal information would motivate them to answer the questionnaires. Therefore, we hypothesized that feedback groups would respond more to questionnaires for which they have visualizations. However, our data did not support this hypothesis; the *A+I* group, which received visualizations only about interruptibility, had the same compliance rates for *Interruptibility* questionnaires and *Mood* questionnaires. This was also true for the *A+M* group, which only received visualizations about mood.

There are two potential explanations. First, while users gained value from the feedback, they did not (or did not bother to) discern between the two questionnaires. Second, the feedback may have more complicated effects than what we envisioned. The feedback not only offered information to the participants but also may allow them to feel more engaged with the study and feel that their contribution is valuable. That higher level of engagement can lead to an increase in overall compliance rate, regardless of which visualization they saw. Future work is needed to explore these hypotheses.

This result also suggests a potential solution for the problem of reactivity inherent in ESM studies, *i.e.*, that measurement or reporting may affect the behavior being studies [1]. For example, a participant in an ESM study about how often a person thinks about happy thoughts may see visualization of the number of happy thoughts and this might cause the participant to be happier. Our study suggests that adding an interesting "dummy" question (unrelated to the goals of the study) and providing visualizations of it to participants can resolve this problem. Since unrelated questions (*e.g.*, interruptibility and mood) are answered the same amount even when only one of the questions is visualized, visualizations of the dummy question can increase compliance to the ESM study and prevent the reactivity effect of seeing visualizations of the behavior being studied.

Future work is needed to determine why visualizations led *Feedback* groups to comply more. One possibility is that feedback may have sparked curiosity. One *A+M* participant looked at the visualizations "to know how many times I have answered the questions and to know about my mood for a week." Another possibility is that visualizations may have revealed self-reflective trends and patterns: "When I was doing something productive my mood was not high." "Apparently, I am least productive with work (Word, Excel) at night when I use AIM. I never use AIM in the morning, thus much more productive with work." Moreover, it is possible that different visualization types

and styles may affect the usefulness of the feedback, leading to different levels of compliance.

CONCLUSION

We have described a new approach of experience sampling called *ES+feedback* that increases compliance rates by providing feedback in the form of visualizations. Our field study demonstrates that showing feedback increases compliance rates by 23%. Additionally, providing feedback about one kind of data from all that is collected in the study is sufficient to increase compliance for all data, thus development time is not significantly increased. We believe *ES+feedback* can be used to improve participation in evaluations of ubiquitous computing technologies. We plan to further explore the advantages and disadvantages of the *ES+feedback* approach and to develop tools to facilitate the deployment of such studies.

REFERENCES

1. Barrett, L.F., and Barrett, D.J. "An Introduction to Computerized Experience Sampling in Psychology," *Social Science Computer Rev.*, vol. 19, no. 2, Summer 2001, pp. 175–185.
2. Carter, S., Mankoff, J., and Heer, J. Momento: Support for Situated Ubicomp Experimentation. In *Proc. CHI 2007*, pp. 125-134.
3. Consolvo, S. and Walker, M. Using the Experience Sampling Method to Evaluate Ubicomp Applications. *Pervasive Computing* 2003, pp. 24-31.
4. Froehlich, J., Chen, M.Y., Consolvo, S., Harrison, B., and, Landay, J.A. MyExperience: A System for In situ Tracing and Capturing of Feedback on Mobile Phones. *MobiSys '07*. pp. 57-70.
5. Hudson, J.M., Christensen, J., Kellogg, W.A., and Erickson, T. I'd be Overwhelmed, But It's Just One More Thing To Do: Availability and Interruptibility in Research Management. In *Proc. CHI 2002*, pp. 97-104.
6. Intille, S.S., Tapia, E.M., Rondoni, J., Beaudin, J. Kukla, C., Agarwal, S., Bao, L., and Larson, K. Tools for Studying Behavior and Technology in Natural Settings. *Ubicomp 2003*, pp. 157-174.
7. Krishnan, A. and Jones, S. TimeSpace: activity-based temporal visualization of personal information spaces. *Personal and Ubiquitous Computing 2005*, pp. 46-65.
8. Palen, L., and Salzman, M. Voice-mail diary studies for naturalistic data capture under mobile conditions. *CSCW 2002*, pp. 87-95.
9. Scollon, C., Kim-Prieto, C., and Diener, E. Experience Sampling: Promises and Pitfalls, Strengths and Weaknesses. *Journal of Happiness Studies 4*, 2003, pp. 5-34.
10. Seligman, C. and Darley, J.M. Feedback as a Means of Decreasing Residential Energy Consumption. *Applied Psychology*, vol. 62(4), 1977, pp. 363-368.
11. Tufte, E.R. *Visual Explanations: Images and Quantities, Evidence and Narrative*. Graphics Press. 1997.