Understanding Contextual Factors in Location-aware Multimedia Messaging

Abdallah El Ali
ISLA - UvA
Science Park 904, 1098 XH
Amsterdam, The Netherlands
lastname@uva.nl

Frank Nack
ISLA - UvA
Science Park 904, 1098 XH
Amsterdam, The Netherlands
lastname@uva.nl

Lynda Hardman
CWI/ISLA - UvA
Science Park 123, 1098 XG
Amsterdam, The Netherlands
first.lastname@cwi.nl

ABSTRACT

Location-aware messages left by people can make visible some aspects of their everyday experiences at a location. To understand the contextual factors surrounding how users produce and consume location-aware multimedia messaging (LMM), we use an experience-centered framework that makes explicit the different aspects of an experience. Using this framework, we conducted an exploratory, diary study aimed at eliciting implications for the study and design of LMM systems. In an earlier pilot study, we found that subjects did not have enough time to fully capture their everyday experiences using an LMM prototype, which led us to conduct a longer study using a multimodal diary method. The diary study data (verified for reliability using a categorization task) provided a closer look at the different aspects (spatiotemporal, social, affective, and cognitive) of people's experience. From the data, we derive three main findings (predominant LMM domains and tasks, capturing experience vs. experience of capture, context-dependent personalization) to inform the study and design of future LMM systems.

Categories and Subject Descriptors

H.5.2 [**User Interfaces**]: Evaluation/methodology, User-centered design, Prototyping; H.5.1 [**Multimedia Information Systems**]: Artificial, augmented, and virtual realities

General Terms

Human Factors

Keywords

Location-aware multimedia messaging (LMM), contextual factors, experience-centered framework, context-aware systems

1. INTRODUCTION & MOTIVATION

Research in Ubiquitous Computing promises to populate our daily lives with specialized 'context-aware' services that enhance our experience of the world by making interaction with it easier, friendlier,

and more efficient [19]. This effort is made possible through embedding (sometimes personal and imperceptible) devices and sensors in our everyday environments. A major advance in this direction is the widespread low-cost availability and adoption of location-aware (or location-based) technologies such as sensor-enabled mobile devices and automotive GPS. Yet despite this unbridled adoption, there is still much to be known about context [8], and how that feeds into our everyday experiences. Inference and adaptation to human intent in context-aware systems is at best an approximation of real human and social intentions of people [2], which requires further exploration of the kinds of services and usability issues brought forth under real-world usage contexts. In short, we need to investigate not only that extent to which context-aware systems need to know about locations, but also people's experiences and their relationship(s) to the location they took place at.

For the present paper, we identify through an exploratory approach the contextual factors surrounding the production and consumption of location-aware multimedia messages (LMMs) with the aim of eliciting implications for the study and design of future LMM systems. Examples of these multimedia messages (MMs) include geo-tagged photos, text, video, audio. These LMMs are anchored to a location by some person, and can be perceived and interpreted by recipients by being at (approximately) the same place where the message was made. Given that locations within cities are rich sources of "historically and culturally situated practices and flows" [20, p. 43], it is reasonable to assume that LMMs can reflect culturally entrenched aspects of people's experiences and make them visible at locations. To this end, use an experience-centered framework to allow us to identify contextual factors that are potentially relevant to LMM.

How do users understand and make use of these LMM systems? What usability issues do emerging LMM technologies give rise to? Is our experience-centered framework suitable for addressing these usability issues? In short, what are the relevant real-world contextual factors involved in creating multimedia messages at locations, and how can these inform the study and design of future context-aware LMM systems? This paper has two main aims: a) To investigate the contextual factors involved in LMM production under an experience-centered framework b) to draw implications for studying and designing future LMM systems. To do this, we adopt an exploratory approach, one that is amenable to the subjective nuances of everyday human cognition and affect.

For our experience-centered framework, we distinguished between two aspects of an experience: process and memory. An experience process (cf., [11]) is a sensory and perceptual process that some person undergoes (through direct participation or observation of events and situations) that results in a change in that person. Given the high variability in computationally modeling and pre-

dicting the process of an experience, for this paper we look mainly at the memory of an experience. Based on the definition of episodic memory given in [18], we define an experience memory as the result of an experiential process, which can be manipulated and actively recalled. It consists of one or more actors, and spatiotemporal, social, cognitive, and affective aspects. It is these aspects of an experience memory that we use as a basis for studying LMM. To this end, we make a contribution towards identifying which contextual factors are important in studying LMM systems, and what kind of experience-enhancing mechanisms need to be supported.

The rest of this paper is structured as follows: first, we review related work on location-aware messaging (Sec. 2). Next, we describe our LMM prototype and the lessons learnt in using it in a pilot study (Sec. 3). Then, we describe our multi-modal diary study (Sec. 4), the category attribution task (Sec. 5) that was necessary for analyzing the diary results, and discuss the assimilated results (Sec. 6). Finally, we conclude by drawing three implications for the study and design of LMM systems (Sec. 7) followed by conclusions and future work (Sec. 8).

2. RELATED WORK

Previous work has focused primarily on location-aware systems that allow users to leave textual messages such as reminders or post-it notes at locations [3, 9, 13, 16]. While these systems support only text, GeoMedia [12] permits attaching multimedia messages (as images, audio or video) to locations. The GeoMedia system however lacked a thorough user evaluation, leaving a gap to be addressed in the study of LMMs, and how they relate to experiences in mobile and ubiquitous environments.

The Place-its system [16] was designed to study how location-aware reminders are used throughout a person's day, the relative importance of locations for reminders, and the effects of reminder message positional accuracy on the reminding process. While reminders may serve as triggers for experiences, the scope is rather narrow. The ActiveCampus application [9] provided insights into how people living on a campus would use such location-aware messages, however, the restriction to a textual medium and an academic surrounding is insufficient for understanding the range of human experiences in everyday settings.

Both GeoNotes [13] and E-graffiti [3] were extensively studied in real-world usage contexts. Studying each provided insight into how people conceived of location-aware systems, the perceived usability of their location-aware functionality, and the relationship between an information and physical space. As in E-graffiti, we are also less interested in tackling the technical problems of context detection, but rather to focus more on evaluating user reception of a location-aware messaging system. Specifically, we want to focus on interesting and novel uses of such a system, and how that can enrich the human experience of being at a media-rich location. However, whereas GeoNotes and E-graffiti were existing application prototypes which were committed to certain design decisions (e.g., in GeoNotes commenting within a note or content-searching using a word-based search engine), we are more interested in the human perceptual conditions involved in LMM with sufficient flexibility to avoid commitment to any one design.

Put differently, our work differs in that we are interested in users' perception of how such systems should be or look like, and not in their reaction to committed design rationales. For example, in Geonotes, the connection between a space and a note was defined explicitly using place-labels, while for us we wanted our users to inform us about the causal relationship between media messages and the entities in a space. Also, while GeoNotes committed to certain types of metadata, we are interested in seeing what kind of meta-
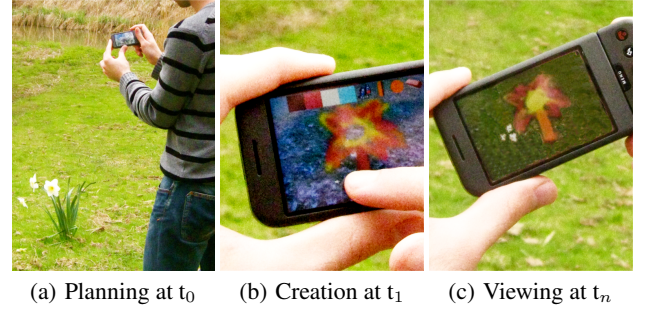

(a) Planning at t_0 (b) Creation at t_1 (c) Viewing at t_n

Figure 1: Interaction with the prototype.

data people would firstly fill in and then later desire to consume. Finally, we wished to study multimodal capture behavior that made use of various types of media (including but not limited to videos, songs, images), and not only location-aware text messages.

3. PILOT STUDY

To understand the experiential factors surrounding LMM, we took a developed prototype application that allows the annotation of locations using three different media types (text, drawing, and photos). The prototype was pilot-tested with 4 subjects where an *in situ* interview method [6] was used to observe experience capture behavior. By annotating locations, the prototype allows users to capture their experiences, i.e., create a digital memory of an experience (Fig. 1(a)). The generated message remains anchored to the location it was created at for later viewing by anyone who has the application installed on their multimedia-enabled mobile device and is at the same place where the message was created.

LMM Prototype

Generation: The prototype application was installed on the Android Dev Phone 1. The initial screen consists of three functions: Create, Snap, and Explore. In Create, a user can create a free drawing (Fig. 1(b)) using touch-based input or type text using the device's keyboard. Here, the location and orientation of the device is retrieved and the user is presented with a camera-view where she can choose to draw or write something. In choosing either option, a snapshot of the camera view is subsequently used as a background canvas for the user to draw or write on. Once a user is finished, the annotated image can be saved. In Snap, a user is taken directly to a camera-view where she can snap a photograph.

Presentation: To view a message, a user has to be at the right position and orientation. In switching to Explore mode, a user is presented with a camera-view, where she is guided to a message by leading her to the creator's original position and orientation. An arrow is drawn on the screen to guide the user towards a message. To indicate the distance between the user's current position and that of the message, the color of the arrow changes within 200m of the message location. Once at the right position, the user can adjust her orientation by looking at a small green indicator arrow shown on the right or left edge of the screen. In doing so, the selected media message is overlaid on top of the camera-view (Fig. 1(c)).

Lessons Learned

While the approach of using a developed prototype provided direct user-feedback on experience capture, all the tested subjects expressed that they had insufficient time to satisfactorily express themselves. Moreover, since the prototype was at its early design stages, users, in capturing their experiences using the provided media forms (drawings, text, photos), were limited by the presented

technology. This created an 'experimental straw man[1]', where it was now unclear what kind of experience-eliciting behavior was being measured: did the users feel that their created LMMs were intrinsically tied to the existing functionality and interaction methods offered by the prototype application, or did they understand that the application was merely a probe into informed user-centric development of future context-aware LMM technology? These concerns are not new: previous work has addressed possible confounds in using location-aware messaging technology in its earlier stages (such as short battery life of the PDAs used in [9] or the sluggishness and effort required for carrying laptops to make messages in [13]). The limitations encountered in previous work and the problems that surfaced in the pilot study led us to revise the chosen method in favor of one that allows understanding LMM behavior for a longer duration and without predisposing users to the functionality and interaction modes of existing technology.

4. MULTI-MODAL DIARY STUDY

The lessons learnt from the pilot study resulted in a redesign of the investigation method. To alleviate the pilot study limitations, we set up a longitudinal multi-modal diary study [1, 17] in order to investigate the contextual factors surrounding LMM production and consumption.

Participants

Eight subjects (6 male, 2 female) aged between 13-27 (M= 23; SD= 4.4) were recruited for the diary study. All subjects were in their 20's, except for S6 who was 13 years old. The reason behind recruiting a young subject was to accommodate a different attitude to technology. Five of the subjects had completed their bachelor's studies, one her master's studies, one pre-master's studies, and finally S6 had completed the first year of high-school. Three of the subjects owned a smart mobile device. All however were familiar with viewing multimedia on such devices and GPS usage. All but S5 declared themselves as social, outgoing people.

Q1	Where are you right now?
Q2	Please explain why you made the media message at this place.
Q3	Please describe how you are feeling right now. (e.g., happy, sad, anxious, excited, lazy)
Q4	Please describe the environment around you.
Q5	Who are you with right now?
Q6	What were you doing before you made the media message?
Q7	Is there an event going on where you are (e.g., sunset, festival, live band, market, dinner)? If yes, please describe the event.
Q8	If yes to question 7, are you participating in this event, or did you only observe it?
Q9	If yes to question 7, is this the first time you participate/observe such an event?
Q10	Were you able to express what you wanted? If not, please state why you couldn't.
Q11	Was there something specific in the environment that you directed this message at? If yes, please state what it is.

Table 1: The second set of questions asked in the diary that pertain to the subject and her context.

---

[1] A straw man is a reasoning fallacy that occurs when an opponent's position is misrepresented. To attack a straw man is in fact to create an illusion of having refuted a given proposition by attacking a superficially similar proposition (the straw man). For us, we adapt the straw man notion to describe misplaced measurement of something superficially similar to what actually should be measured.

Materials

Materials consisted of an information brochure, 8 custom-designed paper diaries, and a set of post-study interview questions. The diaries were custom-designed so that the diary each subject had to carry looked professional and hence would make subjects take the study more seriously, in addition to ensuring that study questions were available for easy look-up. The diary included 2 pages of instructions and 2 pages that contained the 'question template': a set of questions that each subject had to answer after making a message. The question template was split into two parts: questions about the message made and a set of questions about the subject and her context. The first set of questions were: date, time, message media type (drawing, text, photo, video, audio recording, other), title of message, and whether the message is public or private. The message questions (see Table 1) were about: spatiotemporal aspects (Q1, Q4), social aspects (Q5), affective (Q3) and cognitive aspects (Q2, Q7, Q8, Q9, Q10, Q11). The interview consisted of the following questions: difficulty faced in filling in the diary, inspiring days and locations, media preference, environment awareness and overall experience of the past week, willingness to use a future context-aware LMM application, desire to view and write message metadata, and further subject additions.

Procedure

After reading the information brochure, subjects were asked to fill in a personal information form along with a permission statement that permits the analysis and usage of their data. Afterwards, each subject was given a short demo of the LMM prototype, and asked to make two messages with it. This was done as a cautionary measure (as highlighted in [3]) to ensure that subjects understood what was meant by location-aware functionality. Each subject was given a personal diary and an oral explanation about the requirements of the study. Subjects were required to carry the diary with them for approximately one week. They were asked to make a MM (photo, video, text, drawing, song or audio recording) twice per day, so that by the end of the week they had a total of 14 messages. Given the stringent nature of filling in the diary twice per day, subjects were told that they are allowed to make 3 messages per day if they so desired, at the cost of a message on another day. This was done in order to make the testing conditions as natural as possible, under the assumption that there are days where one is more inspired to make messages than others.

The messages made by subjects were restricted to public places, loosely defined as any place outside of their own homes. Upon making a message, subjects were asked, if possible, to immediately answer the questions provided in the 'question template' in the diary. Since subjects may not possess the necessary media capturing device at the time of making a message (e.g., a video camera), they were asked to instead provide an image-based or textual description as a surrogate for the actual message (e.g., a textual description or series of images depicting what a subject's video shot would have captured). At the end of the study, subjects were asked to provide the actual MM either by e-mail or directly through a USB flash drive, return the diary, and sit through a ~10 min. interview. Each interview was captured by means of a tripod-anchored digital camera. After the interview, as motivational measure, each subject was awarded a €20 note and thanked for their participation.

5. CATEGORIZATION TASK

The diary study resulted in 110 user-generated messages, where the interpretation of these was subjective. To understand the motivations offered behind the made multimedia messages, we cate-

gorized subjects' motivations into domain (to what domain does a given location-aware message belong; e.g., entertainment, architecture) and task (for what purpose or task was the message created for; e.g., appreciation, criticism) categories. To ensure the domain and task categories we chose reliably group subjects' message motivations, we needed to account for inter-coder reliability. Therefore, we set up a secondary categorization task that required subjects (distinct from the participants tested in the pilot and diary study) to categorize the motivation responses provided by the diary-study subjects. In order to decide on the best approximate categorization, a voting "winner-takes-all" procedure was applied where a message-classifying category with the most votes wins.

Participants

Six participants (3 male, 3 female) aged between 24-29 (*M*= 26; *SD*= 2) were recruited for the category attribution task. All subjects had completed their bachelor's studies.

Materials

The materials for the category attribution task were the 110 message motivations (i.e., why subjects chose to make the message at a given place) and their corresponding media type, made by the 8 diary-study subjects .

Procedure

Subjects were contacted through e-mail, where they were provided with 110 message motivations and their corresponding media type. They were asked to categorize each message under both domain and task categories (Fig. 2 & Fig. 3), where multiple categories can classify a message. The first set of domain and task categories identified were used as exemplars for subsequent classification. However, if an exemplar category did not suitably classify a given message, subjects were allowed to create new categories as needed.

6. RESULTS & DISCUSSION

The diary-study proved to be a powerful low-fidelity mechanism for studying LMM in real-world contexts without the intrinsic bias evoked from using an existing yet incomplete technology. From the 8 subjects, 2 of them completed only 13 messages, which resulted in a total of 110 MMs. The results of the categorization task provided the basis for further analyzing the diary study data, where the categorization task results were directly assimilated into the diary study results. An equal number of responses to two distinct categories resulted in classifying the message as belonging to both. Below, we present and discuss subjects' media preferences, the identified domain and task categories, the difference between captured experiences and the experience of capture, the different aspects of captured experiences (using out experience-centered framework), and the relevant post-study interview responses given by subjects.

Media Preferences for MMs

To identify what media types should be supported in LMM tools, subjects were asked about their media preferences. From the 110 messages, the most prevalent media types were: photos (45%), text (24%), and songs (13%). The other media types (namely, video and audio recording), were each less than 10% of the total messages made. The lack of video recordings could have been due to the non-availability of the media capture device (e.g., handheld video camera). Only one subject made use of multiple media in a given message, namely 'photo + text' pairs. Not surprisingly, the most chosen media type was photos, which require little cognitive effort to make. For photos made at locations, they can only give
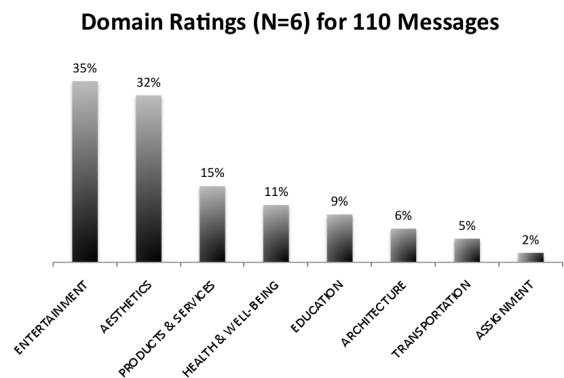

Figure 2: Distribution of domain categories (total = 114%) rated by subjects (N=6) for 110 messages.

a unique perspective on the location, given the high iconic correspondence between a photograph of something at a location and the location itself. As one subject stated when asked about his media preferences "In the beginning, it was photos, and during the week, because it wasn't that interesting, I used more text." Indeed, if the location is not interesting or does not offer any unique perspectives to share with others, then a symbolic medium such as text can be used to express something beyond the qualities of the location itself. Also of interest is S2's remark on using songs because places can remind one of songs, but also because songs themselves can become surrogates for the memory of a place.

Identified and Rated LMM Domains and Tasks

From the initial set of identified domain categories, only 4 out of the 110 messages were problematic to classify. Upon closer inspection, the reason was due to messages where subjects saw it as a duty to make a message (e.g., "Because I had to"). This led us to create an extra 'noise' category: Assignment. Indeed, such problems with subject motivation are sometimes unavoidable during requested study participation [3]. The highest density of messages fell into the Entertainment (35%) and Aesthetics (32%) domains (Fig. 2). Here, aesthetics was defined as something that offers sensori-emotional value (e.g., a beautiful scene), whereas entertainment something that offers amusement (e.g., a film). Only 17% of aesthetic messages were also classified as belonging to the entertainment domain, indicating that there is indeed a distinction to be made. Products & Services (15%) and Health & Well-being (11%), comprising around a third of total messages, are also domains typical of everyday experiences. Overall, the majority of the messages were about the entertainment and aesthetic domains.

Coincidentally, only 4 out of the 110 messages were difficult to classify into task categories. Here, the divergence was mainly between classifying messages as belonging to Appreciation or Criticism. For example, the message motivation made by S8: "It looks sad with the snow" can indeed be understood as both an appreciation statement and subsequent criticism of the state of affairs. Most messages were classified into the Activity-reporting (38%) and Appreciation (36%) task categories (Fig. 3). Activity reporting[2] (i.e., reporting to people what you did) and Appreciation (i.e., enjoying the qualities of something) paralleled mostly the classification of messages into Aesthetic and Entertainment categories, where Ac-

---

[2]Activity-reporting is broader than citizen journalism, which describes activities that pertain to nation- or worldwide events, and not necessarily personal events.
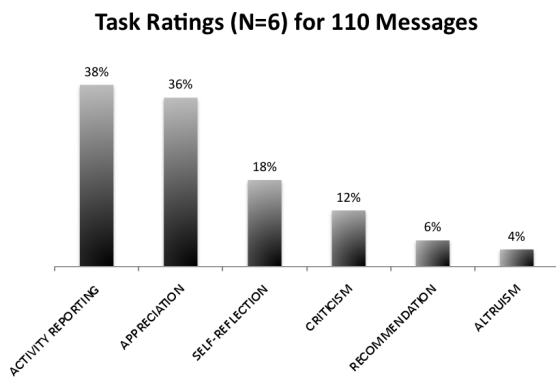
Task Ratings (N=6) for 110 Messages

Figure 3: Distribution of task categories (total = 113%) rated by subjects (N=6) for 110 messages.

tivity reporting messages fell into Entertainment, and Appreciation messages into Aesthetics. Self-reflection (i.e., reflecting on one's own actions or feelings) (18%) indicated that reflecting on one's self is also typical of everyday experiences that warrant capturing. Overall, the findings show that activity-reporting and appreciation were the most prevalent task categories.

Captured Experiences vs. Experience of Capture

When comparing the analyzed data with our subjects' experience with the diary during the post-study interview, the importance in distinguishing between captured experiences (i.e., experience memory) and the experience of capture itself (i.e., experience process) became clear. Whereas captured experiences are information 'about' an experience (*cf.,* the answers to the diary questions), the experience itself is a process emergent from an undertaken activity (*cf.,* the actual experience of using media capture devices and filling in diaries to capture experiences)). For capturing experiences, the aim is to provide an adequate *representation* of a real-world experience that took place (e.g., a community-rated image-based experience sample of a person who parked her bike to photo capture something special in the surrounding scenery). For the experience process however, the aim is to subject users to conditions in everyday settings that would strongly correlate to (if not cause) a desired type of experience while *interacting* with a system (e.g., equipping an LMM system with an adaptive notification system that learns never to interrupt users about new LMMs while driving vehicles).

For the latter, the concern is less about what context is needed to sufficiently re-contextualize the experience of others, but instead about the scoped interaction between the user(s) and the system, where the user experience takes place during the interaction process itself. For the experience process then, we feel the emphasis should be on modeling the user and anticipated interaction with the system. This requires accounting for not only (context-dependent) multimodal input and output support [5], but also the extent the system can make sound predictions about a user's current state to sustain and enhance the flow of interaction [10]. For example, captioning a LMM such as a photo through textual input might interrupt the user's current experience, whereas a voice command label that achieves the same function may occasion a more seamless interaction experience. Additionally, for notification, the system would need to temporally adapt to when users would be most receptive to receiving LMMs, so that the notified LMM can intersect itself gracefully between the user's cognitive and digital life.

Captured Experiences

To understand the different facets of LMM, the results of the diary questions were clustered according to the different aspects of an experience. These are discussed in detail below.

Spatiotemporal Aspects:
For the spatial aspects, subjects were asked about where they were when they made a message (Q1, Q4), giving an indication about their experience at a place. This resulted in the following grouping: Urban (39%), an outdoor setting in the city, such as being on the street; Public Place (21%), an indoor public place such as a café or bar; University/School (17%); Nature (7%), being at a park or nature reserve; Friends/Family Home (6%), at the home of a friend or family member; Home (6%); Transport Vehicle (3%), inside a transportation vehicle such as a tram or metro. Most messages were made in an urban setting, public place, or at the university[3], providing an indication as to the kinds of places future experience-capture technology would be used in. Also, despite that subjects were asked to make messages outside of their homes, a few did not comply, which were classified under the Home category.

For domain and task dependencies in an urban setting, most urban messages fell into the aesthetics domain category (62.9%) and appreciation task category (49%), which highlights the tight correspondence between being outdoors and aesthetic appreciation. Not surprisingly, when controlling for a university/school setting, many of the messages fell into the Entertainment domain category (42%) and Activity Reporting task category (53%), which shows that using such a technology in an academic setting does not necessarily pertain to education. Finally, many of the messages were about Activity Reporting (39%) when controlling for Public Place, which is reminiscent of micro-blogging behavior (e.g., Twitter[4] feeds).

With respect to the spatiotemporal aspects, we were only interested in whether certain days affected subjects' LMM behavior, and not in specific dates and times. For subjects' behavior, S3 and S5 almost exclusively made messages in an urban outdoor environment (78.6% and 71.4%, respectively). Curiously enough, when these subjects were later interviewed about whether there were more inspiring days (temporal dimension) or locations (spatial dimension) in making a message, they reported the following: S3: "Yes, not a particular day, but of an inspiring moment [asked about location] I wouldn't say it was because of the location, it was a matter of coincidence"; S5: "Yes, definitely the weekend [asked about location]; yes, I found that I like changes in my everyday routine places, and when I encounter something that I like a lot that's changed, that's something that inspires me but doesn't happen everyday." From the 8 subjects, 3 of them stated that the location did not provide a source of inspiration, but rather it was coincidental inspiration. However, for inspiring days, all subjects agreed that events provided a source of inspiration, where events included their weekend activities, such as going out for a drink. Overall, these findings are consistent with [16] and [3], who found that the location, in and of itself, is perhaps not an essential part of context, though certainly useful as a trigger for an experience.

Social Aspects:
Subjects were asked about whether they wanted their MMs to be public (visible to anyone at approximately the same location it was made) or private (viewable to only specified networks) [3], as well

---

[3] While arguably a university/school is a public place, the distinction was made here to highlight possible differences between making a message in a non-academic setting and an academic one.
[4] http://www.twitter.com; last retrieved: 28-07-2010

as who they were with at the time of making the message (Q5). Most messages were made public (71%) and the rest private (29%). In analyzing who a subject was with, we defined a person as a single friend or family member and a group as a collection of friends or family members[5]. Nearly half of the messages were made while a subject was alone (46%), compared to being with a group (30%) and with a single other person (25%). However, this might reflect a subject's personality or age; for example S3 made all but one message when alone, and S4 and S8 made more than half of their messages alone (57%). By contrast S6, the 13 year old subject, made most messages while in a group (64%) – this may be because at a younger age, a teenager is usually surrounded by people at home and at school. In considering the domain and task categories for messages made alone, the highest percentage was for the Aesthetics domain (36%), and the highest percentages for the Appreciation (34%) and Activity-reporting (34%) tasks.

The foregoing results illustrate the difference between public and private messages, and messages made alone or with others. While alone-messages dominated our findings, most of these messages (76%) were nevertheless made public. This is in contrast to the findings of [3], who found that notes posted voluntarily were mostly made private. This may be due to their misleading conceptual model that resulted in users treating the E-graffiti system as a limited e-mail system, where E-mail messages are generally addressed to a few private individuals. In assessing the dependencies between the social and spatial aspects, it was interesting to see that half of the messages made alone were made at an urban outdoor setting (50%), which also comprised half of the total number of song messages made (50%). While this may have been a coincidence, it is also not unlikely that when walking outdoors, subjects still feel the need to record their experience, even if alone (e.g., S3's motivation: "To have a memory of this special location").

Affective and Cognitive Aspects:
The mood responses (Q3) of subjects were classified according to valence (positive, negative, neutral, ambivalent) and arousal (high, moderate, low), in accordance with the circumplex model of emotion [14]. We used this model as an instrument for easy and relevant classification of subject's responses according to the valence and arousal dimensions. Most messages were made when subjects were in a positive mood (46%) or highly aroused (46%), where only around half overlapped between these two factors (54%). Negatively valenced (32%) and low-arousal (33%) affective contexts were also prevalent in subjects' responses, compared to neutrally valenced moods (16%) and ambivalently valenced moods (8%) on the one hand, and moderate arousal levels (22%) on the other. It was interesting to see a tendency between being alone and being in a negatively valenced mood (60%), whereas from all messages made in a group, most tended to be positive (55%). For the overlap between negatively valenced moods and being alone, the diary may have functioned as a cathartic outlet for them to express their negative mood, which is also typical of web 2.0 social behavior [4]. This is further supported by the observation that most negatively valenced moods (74%) resulted in messages that were made public. Together, these findings highlight the variability in mood states in everyday contexts, which do not exhibit strong overlap between the location of the experience and the MM.

With respect to the cognitive aspects (Q2, Q7, Q8, Q9, Q10), we were mainly concerned with the causal trigger of a subject's experience (i.e., what in the perceptual environment captured a subject's attention and thereafter served as a trigger to make a message).

[5]Strangers also counted, but there had to be at least one friend or family member for group classification.

Moreover, we wanted to investigate the causal dependencies, if any, that exist between prior subject activity and the created messages. Most messages did not surface a direct causal relation between prior activity and message creation (65%). Yet when there was such a direct dependency (36%), messages tended to fall into either Activity Reporting (39%) or Appreciation (28%). Related to the causal relation between prior activity and message creation were subjects' responses during the post-study interview about whether or not the diary made them more aware of their daily environment. All subjects reported that indeed it did make them more aware insofar as they had to plan where to make a message. As one subject (S2) reported, the diary, if it were a pervasive mobile tool, would not make a difference in raising awareness if it were embedded in daily life. In contrast, S8 stated that the diary was effective in raising awareness by making him contemplate over the beautiful parts of the city he normally takes for granted. This raises the question of whether continuous cognitive access should be designed in an experience-capture tool, so that deliberate planning behavior becomes the norm. Such a mechanism can serve as a persuasion tool to not only create meaningful MMs (cf., in [15], where they use a game-theoretic approach to study selfish user media contribution behavior for designing user incentive mechanisms), but also to raise perceptual awareness of the daily environment.

For the trigger of a message (Q11), most subjects reported that there was something in the environment they directed their message at (60%), however it was surprising that many said there was nothing they directed their messages at (40%). Closer analysis led to distinguishing between three types of triggers and subsequent message classification: Situation (57%), Object(s) (33%), and Person(s) (10%). Here, a situation was defined as a collection of objects that are a pretext for an event(s) or caused by an event(s). Given this typology, it was assumed that if a subject did not direct her message at something specific in the environment, then the trigger of the message was a situation. The high frequency of situation-triggers is consistent with the findings of [13], who found that situation-related chat outnumbered object-related chat.

Post-study Interview Responses

Viewing and Adding Experience-based Metadata:
During the post-study interview, subjects were asked about what kind of metadata (information similar to that asked in the diary) they would like to see if they were using an application that supported LMM. Afterwards, they were asked about their willingness to add this metadata themselves. 5 subjects reported they would like to view such metadata, specifically to see the following: a person's mood, who that person was with, and the event, if any, that relates to the message. When asked about viewing metadata, S7 said: "Some information might be fun to have, like who a person was with, and what event is happening. I would like a context between the message and an event, because the event might no longer be there, and then you would not know it happened at a location, so then it might not make sense." One subject expressed that he would like such information, but only upon request ("Not at first sight, that would ruin my personal view of their message. But it should be available if wanted...why the message was made, what did the person want to express."). The last two subjects found it unimportant to view metadata other than standard attributes such as names, date, and time; S6 [in response to what metadata s/he would like to see]: "Date and time would be nice to see so you know it's a winter photo, and for the private messages to see the name of the person so I know who it is."

Alongside viewing experiential metadata, we also inquired about subjects' preferred methods of being notified about messages at lo-

cations. After exposure to the diary for around 1 week, it seemed reasonable to assume they can tell us about their notification preferences, despite that the study's focus was not on MM notification. Notification in this context means adaptive filtering of messages to subjects' current situation and interests. All but one subject mentioned they would like the future LMM tool to automatically adapt the presentation of messages to their current situation. Only 2 subjects, S1 and S2, specified explicitly the kind of adaptation they would like: filtering by current mood and by date, respectively. The other 5 did not explicitly specify the type of filter, but stated that adaptivity would be the preferred method of handling the hypothetically large number of messages at locations. Despite that most subjects did not have any clear idea how this would be possible, they mentioned that the application adaptivity should depend on the situation they are in, so that it does not become obtrusive; S6: "If I'm walking, then I'd like to search myself, but if I'm biking, I'd like notification of what there is. For example, great nature photos." This indicated that application adaptivity may be best considered as itself context-dependent. The one subject who did not endorse application adaptivity stated that s/he would like to make queries herself through a search function.

With respect to writing metadata, one subject mentioned s/he would fill this kind of information in (S7), four subjects said it would be too much effort (S1, S2, S3, S5), and three subjects said it is contingent on the situation (S4, S6, S8). The latter case is typified by S8's response: "If it would be of any use to me as a user, let's say I filled in 10 of these experiences, and it would say something about what I would like in particular, that would be a nice application to me, so it all depends on the use." However, most subjects (even the ones who thought it would be too much effort to fill in such information) stated that after some time, to make viewing messages more interesting, would start filling in the metadata. This indicates that the problem of filling in metadata can be partially alleviated if potential users are aware of the consumption benefits provided by the metadata (such as more fully grasping the original experience of the LMM creator).

To take a closer look at subjects' metadata writing behavior, subjects' responses were analyzed syntactically according to word count for two factors: the motivation description length for a created message and the environment description length (see Table 2). These two factors were chosen because they generally require elaborate responses to be contextually meaningful, and therefore are indicative of efforts from subjects to fill in media metadata in general. It is interesting to notice the discrepancy in S3's motivation description length (98%), which is at odds with his later response of finding it takes too much effort to fill in the metadata, especially given his relatively high mean word count scores. Also interesting is S7's high discrepancy across environment description lengths (78%); when asked about filling in metadata, s/he said "It's difficult, but yes probably I would fill it in, actually these are a reasonable number of questions; like tagging who you're with, we do that already." While this kind of analysis gives an indication over subjects' efforts and attitudes towards filling in metadata, it may be difficult to generalize these findings to real application usage.

Potential Application Usage
Finally, during the post-study interview, subjects were asked about what type of functionality and interaction they expect from future LMM tools. One subject drew the analogy between such a future application and the microblogging platform Twitter; S1: "I would compare such a device to Twitter, so if there was a device that can instantly post to Twitter a multimedia message, that would be nice, also might be nice to have it just like a diary, to keep a record of

Subject	Motivation Length			Environment Desc. Length		
	Mean	SD	SD %	Mean	SD	SD %
1	13.5	7.6	57	12.3	7.6	62
2	6.2	2.8	44	9.8	5.1	52
3	9.2	9.0	98	15.1	9.6	64
4	6.4	2.6	40	6.3	2.0	32
5	9.1	5.5	61	8.4	6.2	74
6	6.6	3.7	57	15.0	7.1	47
7	7.9	4.9	61	3.9	3.1	78
8	7.0	1.7	24	2.9	2.1	71
Mean	8.2			9.2		

Table 2: Syntactic mean description lengths across subjects for a) reasons provided for created messages b) description of the environment in which the message was created.

what you've done or what you've seen." This latter part of his statement indicates the potential for LMM applications to behave like life-logging applications such as the Affective Diary [17]. Subject responses tended to cover standard online social network interaction: All subjects stated their preference for 'click and share'-type features, indicating that the easier the application used for sharing, the better. Also, nearly all subjects mentioned they would like to comment on other messages (as in Facebook[6]). Related to this, subjects expect to be able to edit their own messages as well as delete them. One subject (S3) stressed the importance of having an optional expiration date for messages, in reference to messages that concern a temporary problem that will likely be resolved in the near future; S3: "To warn/alert others for glass on the street, which can cut into tires").

When asked whether or not they would actually use an LMM tool if one is available, all but one subject said they would. However, all reported that they would not make as many messages; only when the occasion arises for them to express something worth sharing. The one subject who reported that s/he wouldn't use it explained that s/he spent her whole life in Amsterdam, so if s/he would spend time abroad, then s/he would.

Study Limitations

There were two main problems with the diary study: first, making two messages per day for one week may impose an unnatural demand on subjects. In other words, subjects had to sometimes invest cognitive effort in making messages, resulting in noise (*cf.,* Assignment task category). Related to this, subjects were not always able to immediately answer the diary questions (e.g., snapping a photo while walking outdoors), waiting instead until the next opportune moment to do so (e.g., reaching home). The second problem was the availability of media capture devices. Despite that subjects were told to capture anything they wished, so long as they provided a description of what they wanted to capture, a few subjects mentioned they could not express themselves because they lacked the right media-capture tools (e.g., handheld photo camera).

7. IMPLICATIONS FOR LMM
From the above findings, we derive three main implications for designing and studying LMM systems.

1. Predominant domain and task categories in LMM. It seems that aesthetics and entertainment domain categories and appreciation and activity-reporting task categories predominate experience capture behavior (Sec. 6: Identified and Rated LMM Domains and Tasks). This provides a starting point for tailoring future LMM tools to the right target groups (e.g., park visitors, exhibition goers).

---
[6]http://www.facebook.com; last retrieved: 28-07-2010

2. Capturing experiences versus the experience of capture. The diary results made clear the importance in distinguishing between capturing experiences and the experience of capture (Sec. 6: Capturing Experiences vs. Experience of Capture). While capturing experiences requires a method for annotating locations with the *right* kind of information for later intelligent retrieval, the experience of capture requires catering for a *type of interaction* between the user and the system during LMM behavior.

3. Application personalization is itself context-dependent: When subjects were asked about filtering messages, many expressed they would like messages to be shown in accordance with their current situation (Sec. 6: Viewing & Writing Metadata). More importantly, the tension between self-initiated queries and application adaptivity was itself largely a matter of context. This highlights that future LMM applications should not only account for personalized content, but the personalization itself should learn from and therefore adapt to the user's context.

8. CONCLUSIONS & FUTURE WORK

We have taken preliminary steps towards understanding the contextual factors surrounding LMM behavior and how that relates to capturing and consuming experiences. Using an exploratory approach, we were able to derive implications for the study and design of future LMM systems. The collected data in the diary study hinted at the inherent complexity and multidimensional nature of everyday human experiences, where subjective reports did not always offer patterned clues into how to build technology that can support capturing and communicating experiences. Nevertheless, this complexity provides further support for the importance of studying users and their behavior under real-world contexts.

Future work will address human experiences in the context of parks and exhibition venues (cf., entertainment and aesthetics domains), where our target groups will consist of park visitors, exhibition goers, and city residents. The focus will be less on experience memories through LMM, and instead on the experience processes underlying tasks and activities in entertainment spaces. Specifically, we will investigate how interruptions to interaction 'flow' during such tasks by automated sampling methods (e.g., Experience Sampling Methods [7]) can be minimized. Given the uncontrolled nature of real-world contexts in measuring attentional processes, social media content (e.g., Twitter posts) can be used as catalysts for human activity inference and predictive interruption timing [10]. To this end, intelligent experience sampling would not only offer a broader and deeper understanding of system usability issues under mobile and ubiquitous environments, but also provide a harmonious interactional equilibrium between humans and machines when experiences are sampled under real-world contexts.

Acknowledgments

9. REFERENCES

[1] A. Amin, S. Townsend, J. Ossenbruggen, and L. Hardman. Fancy a drink in canary wharf?: A user study on location-based mobile search. In *INTERACT '09: Proceedings of the 12th IFIP TC 13 International Conference on Human-Computer Interaction*, pages 736–749. Springer-Verlag, 2009.

[2] V. Bellotti and K. Edwards. Intelligibility and accountability: human considerations in context-aware systems. *Hum.-Comput. Interact.*, 16(2):193–212, 2001.

[3] J. Burrell and G. K. Gay. E-graffiti: Evaluating real-world use of a context-aware system. *Interacting with Computers*, 14(4):301–312, 2002.

[4] C. N. Chapman and M. Lahav. International ethnographic observation of social networking sites. In *CHI '08: CHI '08 extended abstracts on Human factors in computing systems*, pages 3123–3128, New York, USA, 2008. ACM.

[5] L. Chittaro. Distinctive aspects of mobile interaction and their implications for the design of multimodal interfaces. *Journal on Multimodal User Interfaces*, 3(3):157–165, 2009.

[6] S. Consolvo, B. L. Harrison, I. E. Smith, M. Y. Chen, K. Everitt, J. Froehlich, and J. A. Landay. Conducting in situ evaluations for and with ubiquitous computing technologies. *Int. J. Hum. Comput. Interaction*, 22(1&2):103–118, 2007.

[7] S. Consolvo and M. Walker. Using the experience sampling method to evaluate ubicomp applications. *IEEE Pervasive Computing*, 2(2):24–31, 2003.

[8] A. K. Dey. Understanding and using context. *Personal & Ubiquitous Computing*, 5(1):4–7, 2001.

[9] W. G. Griswold, P. Shanahan, S. W. Brown, R. Boyer, M. Ratto, and et al. Activecampus – experiments in community-oriented ubiquitous computing. *IEEE Computer*, 37:73–81, 2003.

[10] A. Kapoor and E. Horvitz. Experience sampling for building predictive user models: a comparative study. In *CHI '08: Proceeding of the twenty-sixth annual SIGCHI conference on Human factors in computing systems*, pages 657–666, New York, NY, USA, 2008. ACM.

[11] F. Nack. Capturing experience - a matter of contextualising events. In *Proceedings of the 2003 ACM SIGMM workshop on Experiential telepresence*, 2003.

[12] A. Papliatseyeu and O. Mayora Ibarra. Nailing the reality with GeoMedia: location-aware multimedia tags. In *Proceedings of MobiMedia'08 Conference*, Oulu, Finland, July 2008. ACM.

[13] P. Persson and P. Fagerberg. Geonotes: a real-use study of a public location-aware community system. Technical Report, 2002.

[14] J. A. Russell. A circumplex model of affect. *Journal of Personality and Social Psychology*, 39:1161–1178, 1980.

[15] V. K. Singh, R. Jain, and M. S. Kankanhalli. Motivating contributors in social media networks. In *WSM '09: Proceedings of the first SIGMM workshop on Social media*, pages 11–18, New York, USA, 2009. ACM.

[16] T. Sohn, K. A. Li, G. Lee, I. Smith, J. Scott, and W. G. Griswold. Place-its: A study of location-based reminders on mobile phones. In *Ubicomp*, pages 232–250. Springer, 2005.

[17] A. Ståhl, K. Höök, M. Svensson, A. S. Taylor, and M. Combetto. Experiencing the affective diary. *Personal and Ubiquitous Computing*, 13:365–378, 2009.

[18] E. Tulving. What is episodic memory? *Current Directions in Psychological Science*, pages 67–70, 1993.

[19] M. Weiser. The computer for the 21st century. *Human-computer interaction: toward the year 2000*, pages 933–940, 1995.

[20] A. Williams and P. Dourish. Imagining the city: The cultural dimensions of urban computing. *IEEE Pervasive Computing*, pages 38–43, 2006.