

Looking ahead – How field trials can work in iterative and exploratory design of ubicomp systems

Matthias Korn, Susanne Bødker
Department of Computer Science
Aarhus University, Denmark
{mkorn, bodker}@cs.au.dk

ABSTRACT

We investigate in which forms field trials are a workable model as part of an exploratory design process for sporadic, mobile, non-work settings. A major concern of evaluating ubicomp systems is to study how practices and context of use emerge and develop over time when new technology is introduced. To introduce a sophisticated version of our own prototype in the course of an iterative design process, we conducted a public field trial of the system—a new platform for mobile democratic discussions in municipal planning—that we distributed via the Android Market. However, it turned out to be surprisingly difficult to evaluate our design in a setting that stretches over time, place, and without a preselected set of users. Analyzing our difficulties, we develop a general model for methods studying ubicomp systems. On the basis of this model, we characterize an openly interactive approach to field trials in order to look ahead rather than back.

Author Keywords

Ubiquitous computing, methods, field trials, exploration, iterative design.

ACM Classification Keywords

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

General Terms

Design

INTRODUCTION

The history of field trials is almost as long as the history of HCI methodology itself. While the classic cognitivist approach to HCI was primarily coming from a human factors tradition, Bannon [1] in his paper ‘from human factors to human actors’ argued for the need to understand technology situated in realistic use situations and the need to work with users in design. At the same time, Grudin [10] pointed to the added challenges of designing groupware instead of single user technologies. Among his methodolo-

gical concerns are the issues of critical mass, and that designers’ intuition is even more flawed than usual when addressing collaborative technologies.

Both of these authors and many more saw prototyping a possible means of better understanding the future use situation—for users as well as designers. Bardram [2], however, points to new difficulties regarding prototyping of specific sets of applications (CSCW in his case) due to more complicated use settings. Common to these early experiences with the deployment of prototypes in realistic use settings is that the deployment happened within rather well-established use situations and even more well-established communities of practice.

With the new millennium came a new wave of technological and methodological challenges [5]: The technology became increasingly mobile, use situations moved from work to the rest of human lives, and the idea that technologies were designed and deployed as systems one at a time no longer functioned as a basis for design.

Grudin [11] addressed some of these new challenges of ubicomp, in particular that applications are no longer about the ‘here and now’, meaning that use situations stretch into everywhere and forever. This has consequences for the methods with which we analyze and design ubiquitous technologies, because many of the methods deployed hence far were addressing situations where people act, perhaps together, within quite well-understood settings, time spans, and locations. Ubiquitous technologies are often designed for use situations that are not well understood and in the making. Bødker and Christiansen [6] suggested using prototyping to explore which questions to ask in such emergent settings.

Iterative design and prototyping has dominated our research. Accordingly, we see all designs as part of an iterative design process, where the prototypes, for a period, hold on to design decisions [21] and are vehicles for communication in the project and for users’ hands-on experience [7]. Prototypes accordingly are intermediate outcomes that in various forms capture what we know about the product, the use situation, and the design process. Some of these prototypes are versions of the final product that are in various ways fully functional, while others at the other extreme are experimental and throw-away prototypes formed in materials and software that has little to do with a final product (e.g. mock-ups or paper prototypes).

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

UbiComp’12, Sep 5 – Sep 8, 2012, Pittsburgh, USA.

Copyright 2012 ACM 978-1-4503-1224-0/12/09...\$10.00.

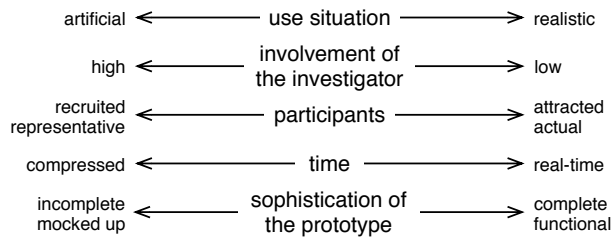


Figure 1: A model of five dimensions for methods studying ubiquitous computing systems.

In general, an iterative design process moves from primitive and ad-hoc prototypes towards something that is increasingly sophisticated and reliable. Yet, there are also movements in the other direction, such as when particular elements of a sophisticated prototype fail and need to be re-iterated in a more exploratory manner. As mentioned above prototypes may be used both to explore settings and to provide answers regarding future designs.

In participatory design, prototypes are explored and reiterated in situations that are remote from the actual use situation (e.g. workshops) as well as situations that are close to real use or located in real use settings. Similarly, use may unfold in real time, or the use situations may be set up as to compress or limit time, such as in many kinds of workshops, e.g. future workshops [16]. The participating users are often selected based on various criteria of representation, but they may also be the actual future users [7]. The prototypes are assessed in settings where the designers are active, and in others where they are more in the background (e.g. think aloud). Figure 1 shows a preliminary model that takes these different dimensions of the use situation, the involvement of the investigator, the participants, time, and the sophistication of the prototype into account.

Upon this background we are challenged by the revocation in ubiquitous computing of more classical scientific approaches. ‘Tests’ and ‘field trials’ have become the new approaches to conceptualizing change, and validating the technologies as such. We argue in this paper that field trials are problematic as part of an iterative design process as they frequently result in *looking back* rather than *ahead*.

Specifically, we analyze a case in a leisure (i.e. non-work) setting, where we have already, in our iterative design process, carried out workshops and walkshops [20] and where we now found it relevant to make a more sophisticated and reliable prototype stand on its own two feet, to see how people would actually use it. In several ways, this case partly failed and we discuss the problems of field trials as part of design based on that failure.

After the related work section, we first very briefly describe the design process as well as the system itself, and second how we went on to conduct a field trial based on a sophisticated version of the prototype. After the description of the study and its outcomes, we analyze this field trial along the

five dimensions shown in Figure 1 in order to motivate a gap in methods for iterative and exploratory design of ubicomp systems.

RELATED WORK

There are numerous considerations as to how one gets about doing field-based evaluation of mobile technologies. Hagen et al. [12] point to various forms of mediated data collection, to simulations and enactments, and to combinations of those as types of methods that are often used in order to assess mobile technologies.

Some of the methods for mediated data collection are in essence clever ways of bringing the classical usability lab out into the field (e.g. [13,19]). Others include various kinds of probing (e.g. [15]) or sampling (e.g. [8]). Other methods yet again are more concerned with having users use the prototype and collecting data through that (e.g. [24,26]), methods that are less concerned with adding additional evaluation layers to the prototype as such. Our study belongs to this latter category. Yet, even though we are logging and gathering use data, we have not added extra evaluation instruments to the prototype. This is primarily because we have been concerned, in this stage of the process, with evaluating and exploring the general concepts of the system, rather than with narrow usability issues that were in focus in earlier parts of the process.

Simulations and enactments belong together with workshops and walkshops to methods where the prototype needs to be less consolidated. Methods that have been applied in this area regarding ubiquitous technologies are, e.g., body-storming [23] and magical things [14]. While we very much believe that such methods are helpful in various stages of iterative design, we have deliberately moved away from simulations and enactments in the work discussed here.

Rather, the intention with this paper is to explore the ways in which we think about field trials as part of a design process. This discussion adds to the renewed interest in method and the value of field trials, and how such experiments should (or should not) be conducted in more naturalistic settings [4,18,22, 25]. In general, we agree with Brown et al. that we need “*much greater innovation in methods around trials, a break away from the assumption that trials should be as ‘natural’ as possible*” [4, p. 1665]. We argue for an approach that does not seek to be ‘natural’ in a mimetic sense, i.e. that does not regard field trials as simulations of use as it would allegedly occur outside and independent of field trials. We instead favor an openly interactive approach with users.

BACKGROUND: A PARTICIPATORY DESIGN PROCESS

In the spirit of the eGov+ project, which explores e-governance services and infrastructure to examine how citizens may be supported in engaging actively in the provision of public services of various sorts, an exploratory participatory design process led to the development of what

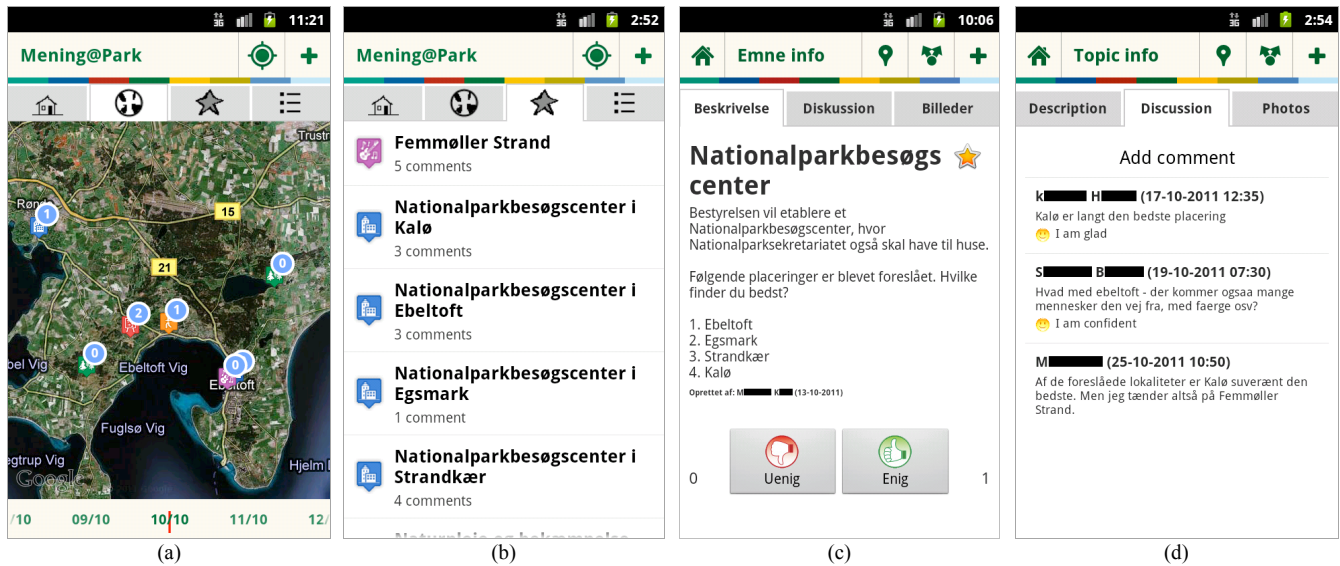


Figure 2. The Mening@Park app allows users, while out and about, to explore topics around the park through location-specific QR codes, location-based notifications, maps, and lists (a, b); to add topics of their own; and to favorite a topic, express agreement or disagreement, comment, and upload photos (c, d).

later came to be called Mobile Democracy. Details of this process and the results can be found in [3]. In the Mobile Democracy case, we worked with a sparsely populated municipality in western Denmark to explore the use of mobile technology in municipal land use planning. The aim was to engage citizen groups in this process by exploring the use of geographical information systems (GIS) and mobile technologies as a means of supporting user involvement. The goal of the overall municipal process is to solicit input and feedback from citizens, among other stakeholders, on a municipal plan that is created every 12 years and continually revised.

Through various design activities with individual citizens, citizen community groups, and municipal planners, we collaboratively explored the design space for potential interventions. We conducted initial interviews with municipal planners and individual citizens as well as focus group interviews with citizen interest groups to get a first grasp of the field and understand potentials for intervention. We held future workshops with citizen community groups and planners respectively to identify concrete problems and opportunities for future developments, and conducted other forms of participatory design workshops for brainstorming and critique.

Throughout the process, we utilized design artifacts such as scenarios, storyboards, and personas, and developed and deployed numerous prototypes of varying sophistication (from paper to fully functional) to better understand the future use situation, for users as well as for designers. With the first more concrete mock-ups and prototypes at hand, we additionally invited citizens for walkthroughs of the prototype and facilitated other forms of hands-on experience during workshops. Specifically, we also conducted what has been coined walkshops [20] to provide hands-on experience for citizens and planners in more realistic use settings. Walkshops do so by urging participants to carry

out concrete tasks on the prototype while out on 30-minute scenario-based walks in a planning area together with the investigators. The primary goal at this earlier stage was not to validate our design concepts or prototypes, but to further explore and understand the design space together with users in more realistic use settings than workshop rooms.

THE PROTOTYPE SYSTEM

One result of this process was the concept and implementation of two interconnected prototypes: a mobile app and a browser-based version that both were designed to facilitate citizens to discuss topics located on a map (and in the world) while out and about. The simple idea behind this concept is to gather feedback and opinions in the moment (when users are physically close to the planning object) and allow for deeper reflection and collaboration later (remote to the planning object) [3]. The rest of this paper mainly concerns the ‘evaluation’ of the mobile app.

The intention of our design is to facilitate discussion and collaboration among citizens and between citizens and municipalities, rather than merely one-way, location-based information from citizens to the municipality. Thus, using the platform, citizens can explore, create, and react to topics created by other citizens or by the municipality in a number of ways: through maps, lists, expressions of agreement, comments, and photos. A special focus in the latest iteration of the mobile app was on facilitating users to establish a connection between a digital topic and a physical place. The app therefore offers various ways to link topics to locations and vice versa: location-based notifications, access to topics through location-specific QR codes, as well as locating a specific topic on the map.

In order to advance our exploration of mobile deliberation systems, the Mobile Democracy app was transferred into a new domain, adapted, and renamed to Mening@Park

(‘mening’ being the Danish word for opinion or point of view). Figure 2 shows screenshots of the final app.

EXPLORING PROTOTYPE USE

With a concept and working prototype at our disposition, we set out to explore, how our design could be introduced into more real-life settings. In this section, we will describe the general domain this took place in, the concrete study context, and the preparations for the study itself.

A new domain

When we moved to a new domain, insights from facilitating participatory land use planning in a municipality were transferred and applied to the participatory planning process of a national park (NP). This case explored the citizen participation process in the forming of Mols Bjerger, the second of five NPs to be established in Denmark. Under the supervision of a board and secretariat, a central aim of this development was a plan for the first six years of the park’s official existence.

The park is somewhat unconventional in that it is not merely a nature reserve, but people actually live and work there. Thus, residents have a strong interest in its development. Similar to municipal planning, many topics regarding the planning process could be connected to specific locations in the park and visualized and discussed on a map. Therefore, deploying the Mobile Democracy prototype in this domain presented a viable option for further exploring our ideas and concepts.

Leading up to the here presented field trial, the development of the NP had already been followed and studied by us for over a year. Well aligned with this development, we conducted exploratory, qualitative interviews with the NP secretariat and participated in public hearings in a first hearing phase. The aim was to establish an in-depth understanding of the citizen involvement process in the NP.

After phase one, a long internal process between the secretariat and the board evaluating all the feedback and suggestions received led to the release of a draft NP plan. In a second public hearing phase citizens had the opportunity to dispute and submit proposals to this draft for about three months. Before the actual start of the second phase, we reconvened with the NP secretariat in two sessions to brainstorm and discuss possibilities for concrete interventions—which we report on in this paper.

In retrospect, our activities have thus not only been motivated by, but have also been deeply integrated in, and aligned with, this ongoing participation process and were fully supported by the NP secretariat as part of their effort to solicit feedback and proposals.

Deliberation in the park

In order to motivate our setup and compare to our own data, we have coded the publicly available comments and proposals that individuals and organizations have made to the

draft NP plan¹—either off-line or by visitors to the tent at the Ebelfestival (see next section)—during the second public hearing phase and in parallel with our own study (see Table 1). Our interests have been two-fold: First of all, how many responses were made by individuals, versus those of NGOs and other organizations (e.g. the municipal council), and therefore of lesser interest to Mening@Park. Secondly, which comments were about particular localities or locations, versus those who were general, and again for that reason not the primary type of contribution possible through Mening@Park.

Submissions	Total	From individuals		From organizations	
		About location	No location	About location	No location
Ebelfestival	4	3	1	0	0
Off-line	55	20	9	6	20

Table 1: Comments to the draft NP plan from individuals and organizations.

The total 29 location-specific comments include, e.g., very specific comments with attached GPS coordinates to suggest park benches, as well as comments that make reference to particular buildings, or geographical positions to indicate ends of suggested walking or biking paths. With more than half of the comments to the NP plan being made by individuals, and more than two third of these being about specific locations, this underlines a general applicability of the Mening@Park concept in this context and a general interest in the development of the park.

Study context

The actual intervention took place as a deployment of the Mening@Park system, which was strategically kicked off at the Ebelfestival, an annual and week-long local folk and harvest fair in Ebeltoft, the biggest city in the park.

The festival attracts many local, domestic and international visitors—officially 40.000 in total in 2011.² It is a busy place for tourism in the park at large (it conveniently takes place during the autumn school holiday week), but more so for local community building and the ongoing citizen participation process with many local inhabitants as well as summerhouse owners visiting the festival. With the second public phase in full swing, the NP set up a dedicated, fairly

¹ See Nationalpark Mols Bjerger (2011), Høringssvar til Forslag til Nationalparkplan 2012-2018 Nationalpark Mols Bjerger. <http://www.danmarksnationalparker.dk/Mols/Organisation/Nationalparkplanen/horingssvar2011.htm> (last accessed on March 20, 2012).

² See Nationalpark Mols Bjerger (2011), Nationalpark Mols-bjerger sender 8250 planer ud. <http://www.danmarksnationalparker.dk/Bibliotek/Mols/planudsendelse.htm> (last accessed on March 20, 2012).



Figure 3: The two study settings during the Ebelfestival: the NP tent on the festival site (left) and the gate (with our QR sign) leading to the Kalø castle ruin in the back (right).

visible tent close to the entrance of the festival site (see Figure 3). At the tent, the NP distributed the draft plan and disseminated 25 concrete action proposals from the plan through roll-ups around the tent. They additionally invited changing collaborators to share the tent space (such as the local hunters association and the natural history museum) in order to offer various hands-on activities for children and attract more visitors.

The NP's main goal was to create public awareness of the NP itself, the NP plan, and its topics by getting into dialogue with people. The NP reported 2.000 visitors to the tent during that week.³ We used the festival and the buzz and activity around it as the main stage for deploying our prototype and studying its use. We did not, however, recruit participants (in the traditional sense) for using the system in general. Before elaborating on the intervention in the next section, we hint at some of the preparations necessary to address the challenges discussed in the introduction.

Preparing the prototype deployment

Leading up to the festival, we not only prepared the Mening@Park prototype for the study purpose, but also polished it for real use by real users in the field. These preparations included the prototype itself, the initial content that was added, as well as other activities to advertise the app to gain critical mass.

Preparing the prototype included branding it for the NP, implementing various features that were only hinted at before but central to our study, removing obvious usability problems and other small bugs identified through heuristics so as to not have them distract from our actual goals, streamline the registration and log-in process for easy access and use, internationalization of the user interface in English and Danish, and, finally, adding interaction logs and other means of documenting actual use for later analysis.

A second central part of the preparations was to add 'real' and meaningful rather than no or mock-up initial content to the system. As [7] noted, this is often important in order to

be able to assess prototypes in the field. In our case, this meant providing easier access to the system and the community by allowing users to contribute first with comments rather than creating topics of their own from the very start. The seven initial topics that we thus added were almost exclusively proposed directly by the NP secretariat rather than by us. They were therefore relevant for the overall process, and the secretariat wanted to stimulate discussions and receive feedback on them. Three of the seven topics consisted of three or four possible geographical locations for placement, which we added separately due to a conceptual limitation in the prototype that tied a topic to only a single location. We thus had a total initial topic count of 15.

We tried to address the challenge of achieving critical mass as highlighted by Grudin [10] by situating (and integrating) our intervention in a lively and meaningful context as outlined above. Our goal was to introduce the prototype as an open platform available for everyone interested rather than a closed user study with a number of recruited participants. We thus made the app available on the Android Market, a wide-distribution channel for mobile phone applications, from where citizens could easily install it to their Android smart phones. The Android Market, however, is merely a convenient distribution channel: it creates new methodological challenges and opportunities, but is not a tool for investigation in itself [22]. We announced and advertised the existence of the app through various means (a pertinent and active Facebook group, emails to interested stakeholder groups and concerned individuals, etc.).

At the festival site, we had a big poster prominently hung up in front of the tent and fliers on the table advertising the app, which we also handed out to visitors pro-actively. During the festival week, and as one focus of our study, we also hung up eight A4-size signs with big and intriguing, location-specific QR codes at the tent and at various popular locations around the park, which attracted attention for some. The QR codes potentially provided direct access leading to a topic about that specific location.

CONDUCTING THE STUDY

The study can be divided into four parts:

1. deployment of Mening@Park at the festival tent,
2. deployment of Mening@Park at the Kalø castle ruin,
3. workshops at the festival site, and
4. post-festival study activities.

The ongoing deployment of the Mening@Park system constitutes the basis for all four parts.

The first three parts took place during the Ebelfestival, where we also first introduced the prototype. Here, we initially split our attention between two different settings: the NP tent at the festival site and the Kalø castle ruin, a very popular destination in the NP (see Figure 3 for an impression of the two settings). The reasoning behind this was to benefit from the generally high and concentrated (recrea-

³ Ibid.

tion) activity in the park during the holiday week by adding a second setting more realistic in the sense of experiencing the park as such. This part comprised of roughly three full days of activities.

Towards the end of the festival week we decided to adapt our strategy and approach to the reality of the festival situation by starting a more concentrated effort to interact with participants by conducting workshops on the final festival day (a Saturday). This interestingly contrasts the ‘uncondensed’ activities during the first part of the week, which we will return to in the discussion. After the festival week, the system was kept running and we conducted various post-study activities.

Deployment of Mening@Park at the festival tent

At the festival tent, we advertised the app, approached people about their mobile phone use, and observed how visitors interacted with the provided information and the NP staff at the tent. Our activities included individual in-situ demonstrations of our system, informal impromptu interviews with visitors and NP staff, as well as observations of activity at the tent. Interviews were audio recorded, observations were documented in field notes and photos, and usage of the app was logged. The focus lay on common and specific information enquiries and questions by visitors as well as the uptake of our app and QR signs at the tent.

Deployment of Mening@Park at the Kalø castle ruin

The Kalø castle ruin, which was our second study setting, is one of the most popular destinations in the NP—for tourists and locals taking a walk alike. We hung up a QR sign on each side of a gate leading from the parking lot to the ruin, and spent time observing people’s use of it. We partly shadowed them to see the extent to which they engaged with the app as well as their general use (and non-use) of mobile technology during the visit.

Workshops at the festival site

For the last full day of the festival, we decided and planned to conduct more concentrated activities in the form of workshops in the area surrounding the festival site [20]. We tried to recruit some local, concerned citizens (including members and representatives of stakeholder organizations in the park) that had interest in the park’s development through Facebook, email, and through the NP staff beforehand to come to the festival and participate in our workshops. However, no one signed up and it was difficult to enlist people for a 20-30 minute walk on the spot. Therefore, instead of bringing people to the sites and artifacts of study, we had to sacrifice some of our intentions of realistic settings in this instance and bring the artifacts to them.

We took two different QR signs that were about topics in the immediate vicinity and pro-actively approached random people around the tent and the whole festival site. We encouraged them to scan the QR codes with provided phones and involved them in a brief try-out of the system combined

with an informal interview. As a vantage point, we specifically enquired about familiarity with and opinions about QR codes as well as mobile technology use in general and in the park. Interviews were again audio recorded and additionally documented in field notes. Altogether, this resulted in about 20 informal interviews with individuals, couples, families, groups of friends, etc. and, while not being the most critical part, provided the most interesting insights for our ongoing exploration of the field.

Post-festival study activities

After the festival, the prototype was maintained and kept running until the end of the second public hearing phase. The system has been further advertised at two consecutive exhibitions at a local library and at the town hall. Lastly, all three people from the NP secretariat attending to the tent during the festival were interviewed about their experiences of getting in direct contact with citizens at the festival in general and the deployment of our prototype in particular.

USAGE STATISTICS

Here we provide a brief overview of usage statistics of the deployed prototype system during the complete field trial period. The numbers were derived from interaction logs, our database, and statistics provided in the Android Market developer dashboard.

The prototype has officially been running until the end of the second hearing phase. This amounts to a runtime of almost two months (52 days or 7,5 weeks). The Android Market reports 46 ‘total installs’ of our app and 27 ‘net install’ at the end of the official runtime (the maximum ‘net installs’ was 33). During the official runtime, 27 users initially registered with the system, 20 (74%) of them during the festival week. That means for 59% of the total installs a user was registered in the system (not accounting for double installs and other irregularities).

Users could contribute to the system in four ways: creating topics, adding comments to topics created by others, expressing agreement or disagreement, and uploading photos. Only one topic has been created by a user, in addition to the 15 we added. Nine of the 16 topics received in total 29 comments (with a maximum of five comments per topic) by eleven users (with a maximum of nine comments per user). 17 (59%) of the comments have been made in the first week (the festival week), and another nine in the week thereafter (in total 90% of all comments). In total 61 agreements or disagreements have been expressed on all 16 topics by 21 users. No photos were uploaded. Six (22%) of the 27 registered users have not contributed to the system at all.

Summarizing, use of our system has been concentrated in the festival week and the week thereafter. Afterwards, the system has only been used very sporadically. We judge the number of users to be fairly decent for the short period of time that the festival had created activity. However, activity

by user was, not unexpectedly with this marginal type of app, quite low.

Comparing to the general deliberation in the park (cf. Table 1), visitors left three relevant topics at the festival tent and the general hearing left 20 topics comparable to ours (left by individual users in relation to a particular location) in addition to six location specific topics left by organizations. We did not have access to these documents until after our study, but we believe that these numbers mean that the idea of commenting on the NP and the park plan as such was highly relevant to a number of people and organizations in this time period. Hence, the reasons why we did not generate more use must be found in the study, rather than the domain. In the following we turn to elements of our study to discuss why.

ANALYSIS

As described in the introduction, we see iterative design processes as a number of activities that can be characterized along a number of individual dimensions. These dimensions relate to the realism of the use situation, the involvement of the investigators, the participating users, the sophistication of the prototype, and time (cf. Figure 1).

What is typically understood as a ubicomp trial use exists at one end of these dimensions where the finished prototype is used by voluntary users in real, un-tampered use situations at a safe distance from designers and researchers. On the other extreme, we typically find mock-up prototypes used in workshop settings in close interaction with the investigators. However, there are many other ways in which these dimensions may be combined and in the following we use the dimensions to analyze and discuss our case.

Use situation

The different settings of our study have different properties in how closely the setting resembles the actual use situation that we envisioned.

The Kalø castle ruin, as a typical location within the park but outside the buzz of the festival, clearly represented a setting that was closer to the envisioned use situation, ‘out there’ in the park where people meet concerns and issues that they want to share with others (and the NP board). At this setting, we tried to make the study have a very withdrawn role, hoping that people would, on their own accord pick up the QR codes and start using the system. At the same time, this meant that we had very few means of ensuring that the system was actually used, as we discuss later.

The NP tent at the festival site provided a more artificial version of the use situation, both because the site as such was created with the purpose of the festival and was not a permanent site of the NP, and because, as a consequence of the festival, the buzz that it created was slightly artificial.

Due to this, however, people were prepared to ask questions, to be informed, and even to be approached by tent

staff on matters of the NP. However, they had no expectations of being approached and used as informants by the investigators on the matter of smart phone technology. While the prototype itself provided a ‘ticket to talk’, even in this setting of citizen interaction, we suspect visitors frequently perceived our interaction with them more as a unidirectional demonstration and advertising situation rather than a bidirectional study situation. Consequently, they often felt intimidated and were limited in their feedback.

Slightly frustrated by this communication mismatch we decided to involve participants more structured and directly. In the actual workshops in the third part, this meant that we tried to use the momentum of the festival while sacrificing some of the realism of the use setting. Due to no one having signed up upfront, we were restrained to walk around the festival site, rather than the surrounding downtown area, where access to real NP plan issues was quite sparse (e.g. a visitors center could not realistically be build on the festival site itself, but rather 200 meter away from it).

Eventually, the tent setting turned out to be very beneficial when talking to the NP staff. They were talking with citizens all day. This meant for us that we could have very interesting and fruitful informal interviews with them about specific encounters with citizens. In that regard, the tent presented a good context for talking to NP staff different from and in addition to the more formal interviews and workshops at their or our location. The tent presented a practical context, where they had the citizen perspectives immediately visualized.

Involvement of the investigator

The different settings of our study, in addition to how closely they resemble the actual use situation, are also different in terms of the involvement of the investigator in the use or study situation.

At the Kalø castle, people were left on their own, completely undisturbed by the investigators. However, it turned out that the possibility of learning what people were doing and thinking (beyond seeing if they knew QR codes) was extensively difficult. With the setting being this ‘wild’, there was a limit to what artifacts people could interact with, what incentives they had to do so, and consequently what could be observed by us. Basically, people also did not tour the castle to be interrupted and questioned by investigators, but rather to enjoy a brisk walk or be with family and friends, on top of seeing the historic remains. In addition to making use infrequent, which in itself posed a challenge e.g. with respect to how it could be observed, it also made it very challenging for the investigators to approach visitors e.g. for interviews. These situations did not lend themselves well to interruptions. Both the use situation and the level of interaction with the investigator were problematic in this setting and, consequently, this setting of little use to us.

As described previously the activities at the tent were entirely different. Even though the situation lent itself more

to interaction and intervention from the investigators and in a way thereby reduced the realism of the use situation, the involvement of the investigators as such was more realistic: it was a fair and people were in general being approached by staff at the tent. Nonetheless, as described, the framing of this setting may have made people see investigators as sales people which was quite problematic to the study.

In an attempt to move away from this, the workshops were set up. We were hoping they would frame the role of the investigators more specifically as researchers. However, whether that ended up being the case, is an open issue since participants were, in the end, enrolled from the crowd at the festival and not based on the distributed invitations. Yet, simply adjusting the level of investigator involvement at the workshops helped dramatically increase the number of insights we could gather. At this time we were not so much concerned anymore with attracting users for the field trial itself, but rather pro-actively interacting with visitors to collect qualitative data for further exploration.

Participants

As mentioned earlier, generating a critical mass of activity is a crucial problem, not primarily with respect to the use of the end product [10], but rather in order to make trial use work as such. For this purpose, many evaluation methods used in ubicomp research frequently either (1) pro-actively and consciously speed things up and ‘make use happen’ by, e.g., introducing artificial tasks for users or setting up ‘use sprints’ (e.g. workshops), (2) introduce artificial incentives to the system that were not part of the original design concept (e.g. some form of gamification), or (3) create artificial incentives for users outside of the system itself to use the system more than they otherwise would (e.g. compensation, provided equipment, commitment through personal relationships [4]).

In an example for (3), Brown et al. [4] talk about trial use in situations where the researchers have enrolled users into the study. In these recruitment situations they found that “*users shape or enhance their behaviour in a trial or experiment, in response to the imagined desires of the investigators*” [p. 1660]. They state “*much of participants’ motivation to use the system seemed to stem from a sense of obligation to us rather than their relationship to the system per se*” [p. 1660]. To avoid, e.g., that users are more likely committed to the investigator rather than the idea, we largely steered away from such artificial incentives by publicly deploying Mening@Park to an ‘open market’ [22] instead of actively recruiting users for the field trial. We wanted not only representative, but actual future users.

However, we had problems getting enough people to use the system during the field trials in order for them to reflect on it. We started with a rather purist idea that it would be beneficial to let users come to us rather than the opposite, and in this sense we would get more authentic use. Along the way we realized that a closer and more real contact than

the announcement posters was needed, hence the interviews at the tent, the emails to relevant stakeholders, and the workshops. In retrospect, we should have worked more along this continuum up from the start. Perhaps by getting more enrolled use corresponding to (1), we would ultimately also have ended up getting more authentic use, simply due to critical mass of comments etc. that would ultimately motivate even more comments. Above all, with a bustling platform we might have eventually also gotten richer insights and perspectives from users.

Time

We largely let field trial use of Mening@Park progress in real time and without our intervention. This, to some extent, resulted in the limited success of our system. Initially, we in fact were working under the assumption that field trials should be conducted as ‘natural’ or ‘real’ as possible.

We were convinced that the Ebefestival should help to generate some usage of our prototype system. Due to the nature of the festival, the activity was nicely compressed time-wise, making it easier for us as investigators to monitor and study. Yet, the buzz surrounding it was not sufficient to gain momentum. One question is if we were simply too impatient? The usage data suggest that even shortly after the Ebefestival, no continued discussions took place. Evidently, we have no way of knowing if some external event would at a later state have led to further discussions.

Hence, for the workshops, we wondered instead what would happen if we accelerated use in order to spur discussion about it: The workshops condensed events to a shorter time frame and ‘made things happen’ that would normally take place over longer periods of time (or not at all). This intensified activity through time compression resulted in broadening the discussions with participants towards the bigger picture, e.g., ranging from how likely they are to use such discussion platforms, to the greater impact such technologies might have on the park’s development.

Time is a well-observed variable in ubicomp field trials. It is a generally acknowledged tension that studying social systems in a rather short time frame is problematic as distinct social practices are not likely to emerge and solidify. Referring to [11], the ‘everywhere and forever’ nature of many ubicomp systems stands in contrast to the comparatively short-lived nature of the deployed research prototype. Accordingly, with our field trial alone we are not able to, by only observing a snapshot in time, make any conclusive comments on how the system is and always will be used. With the investigation discussed in this paper, it is evident that it only embraces the early stages of use-in-the-making, and not the longer-term development of the use activity.

With this perspective in mind, we found that it is difficult to both study use as un-tampered by the designer, and look ahead at the same time. Studies based on trials that ‘leave use alone’ will always look back, and hence cause problems as part of an iterative design process.

Sophistication of the prototype

The level of sophistication our prototype reached and that it could basically stand on its own two feet was a starting point for us, and one reason why we embarked on a field trial in the first place. Our assumption was that this was the next logical step in an iterative design process.

Hence, the polishing of the prototype, the fliers, and the poster were supposed to make it appear as a more or less finished service ‘in beta’. The initial content that we added supported this appearance and hinted at some activity already going on in the system. However, while the prototype was as such self-contained and fully functional, it did not feel completely finished. The way we thus eventually presented the system to festival visitors was rather framed as some experiment by a university, a pure try-out to explore further the possibilities of such applications. This was reflected in the way we, and the NP staff, talked about it, e.g., at the tent: “*a prototype*”, “*an experiment*”, “*not yet a finished system*”, “*yes, that’s how we envision it to be*” etc.

This portrayal of the system as incomplete and preliminary to some extent diminished the usefulness of the prototype as a probe, as a real service proposition. By lightly acknowledging its limitations rather than being bold, we tended to discover and discuss issues with the concrete artifact rather than the broader perspectives and potential impact. For instance, a common theme was that it was too much hassle to install the app just to be able to see or make comments. We knew that and a mobile website with limited functionality was part of our overall concept. Yet, it took over many of the conversations with participants. Therefore, reflecting on the aim of our study, it is important to gauge the sophistication of the prototype system with respect to the purpose it needs to fulfill within the study.

IMPLICATIONS FOR RESEARCH

In critically analyzing where our study is located on the individual dimensions of our model, we were able to identify tensions in the choice and application of our methods. E.g. while our starting point was a sophisticated version of the prototype for people to try out and use on their own, we found that we actually fell short of interacting with participants sufficiently and in a way that would help us to further explore the future application of mobile deliberation technologies in the park. Many ubicomp systems, by virtue of being brought into sporadic and mobile settings, are difficult to study, as use practices are really only in the making [11]. Hence, the application of field trials is problematic. In our case, people do not, already, walk around in nature and comment on a plan, even if they are concerned citizens that otherwise take part in local democracy.

By looking at these dimensions one at the time we see not only a need, but also a path away from ‘naturalistic’ field trials in a conventional sense. As with Brown et al. [4], our experience has led us to think that such ‘naturalistic’ field trials are difficult to set up and may not be as informative as

expected. We have come to think that an openly interactive approach to field trials might prove to be both more viable and more insightful. Pursuing this approach, we attract, rather than recruit, actual future users to an open system. We interact with (some of) them directly and take their experiences of using the system as a springboard for in-depth explorations of their perspectives and insights about the concrete system, the broader concept, and potentials and tensions for future approaches in general.

There are, however, fundamental epistemological issues underlying the challenges of field trials in design. Field trials fundamentally look at use as it is here and now, and possibly as it has developed to become what it is, if the field trial is left to happen long enough. Design on the other hand, is concerned with understanding and shaping the future. Anticipating future use is fundamentally difficult, which is exactly why Floyd [9] and many others have emphasized the need for iterative design, where prototypes are at best versions of a product that may later be changed, once it has been tried out in use. While even such rather sophisticated prototypes need to be tried out and evaluated, they quite fundamentally never have the permanence that justifies a ‘natural’ field trial. Accordingly, classical field trials fundamentally *look back* on the past in that they study what participants have done, an insight that is definitely needed in the long run. At the same time, design needs to work with fully functional prototypes to *look ahead* and explore what participants might do towards use practices that are still emerging and in the making. This is a fundamental methodological challenge that we are facing with the current state of ubicomp technologies and practices.

CONCLUSION

Our starting point of this paper was that we wanted to add ‘naturalistic’ field trials to our design process in a leisure setting. Our experience has led us to think that such ‘naturalistic’ field trials are difficult to set up in these settings and may not be as informative as expected. We thus argue that field trials might be better thought of as a part of, rather than as an assessment of, an iterative design process—i.e., a counterpoint to the more prevalent ‘design then study’ mentality in ubicomp. We have come to think that an openly interactive approach to field trials might prove to be both more viable and more insightful.

In setting up trials or prototype evaluations, we may move back and forth along a number of dimensions where realism of the use situation and sophistication of the prototype are at stake. How we may position a specific trial on these dimensions is a methodological choice. However, this is not an either or. Our model helps to gauge this gap between prototyping and workshops on the one hand and full-blown field trials on the other by locating critical aspects of these methods on the individual dimensions. While classical field trials fundamentally look back on the past, we need ways of working with fully functional prototypes to look ahead towards use practices that are still in the making.

ACKNOWLEDGEMENTS

We wish to thank Nationalpark Mols Bjerger for their support and collaboration as well as Barry Brown and the anonymous reviewers for their comments.

REFERENCES

1. Bannon, L. From human factors to human actors: the role of psychology and human-computer interaction studies in system design. In Greenbaum, J.M. and Kyng, M. (Eds.). *Design at work: cooperative design of computer systems*, Erlbaum, 1991, 25-44.
2. Bardram, J.E. Organisational prototyping: adopting CSCW applications in organisations. *Scandinavian Journal of Information Systems* 8, 1 (1996), 69-88.
3. Bohøj, M., Borchorst, N.G., Bødker, S., Korn, M., and Zander, P.O. Public Deliberation in Municipal Planning: Supporting Action and Reflection with Mobile Technology. In *Proc. C&T 2011*, ACM Press (2011), 88-97.
4. Brown, B., Reeves, S., and Sherwood, S. Into the wild: challenges and opportunities for field trial methods. In *Proc. CHI 2011*, ACM Press (2011), 1657-1666.
5. Bødker, S. When second wave HCI meets third wave challenges. In *Proc. NordiCHI 2006*, ACM Press (2006), 1-8.
6. Bødker, S. and Christiansen, E. Designing for ephemerality and prototypicality. In *Proc. DIS 2004*, ACM Press (2004), 255-260.
7. Bødker, S. and Grønbaek, K. Users and Designers in Mutual Activity—an analysis of cooperative activities in systems design. In Engeström, Y. and Middleton, D. (Eds.). *Cognition and Communication at Work*, Cambridge University Press, 1995, 130-158.
8. Consolvo, S. and Walker, M. Using the Experience Sampling Method to Evaluate Ubicomp Applications. *IEEE Pervasive Computing* 2, 2 (April 2003), 24-31.
9. Floyd, C. Outline of a Paradigm Change in Software Engineering. In Bjercknes, G. Ehn, P. and Kyng M. (Eds.). *Computers and democracy—a Scandinavian challenge*, Aldershot: Avebury, 1987, 191-212.
10. Grudin, J. Groupware and social dynamics: eight challenges for developers. *Commun. ACM* 37, 1 (1994), 92-105.
11. Grudin, J. Group dynamics and ubiquitous computing. *Commun. ACM* 45, 12 (2002), 74-78.
12. Hagen, P., Robertson, T., Kan, M. and Sadler, K. Emerging research methods for understanding mobile technology use. In *Proc. OZCHI 2005*, CHISIG of Australia (2005), 1-10.
13. Høegh, R.T., Kjeldskov, J., Skov, M.B. and Stage, J. A Field Laboratory for Evaluating in Situ. In Lumsden, J. (Ed.). *Handbook of Research on User Interface Design and Evaluation for Mobile Technology*, IGI Global, 2008, 982-996.
14. Iacucci, G., Kuutti, K., and Ranta, M. On the Move with a Magic Thing: Role Playing in Concept Design of Mobile Services and Devices. In *Proc. DIS 2000*, ACM Press (2000), 193-202.
15. Isomursu, M., Kuutti, K. and Väinämö, S. Experience clip: method for user participation and evaluation of mobile concepts. In *Proc. PDC 2004*, ACM Press (2004), 83-92.
16. Jungk, R. and Müllert, N.R. *Zukunftswerkstätten, Mit Phantasie gegen Routine und Resignation*. München, 1989.
17. Kensing, F. and Madsen, K.H. Generating visions: Future workshops and metaphorical design. In Greenbaum, J.M. and Kyng, M. (Eds.). *Design at work: cooperative design of computer systems*, Erlbaum, 1991, 159-168.
18. Kjeldskov, J., Skov, M.B., Als, B.S., and Høegh, R.T. Is It Worth the Hassle? Exploring the Added Value of Evaluating the Usability of Context-Aware Mobile Systems in the Field. In *Proc. MobileHCI 2004*, Springer (2004), 529-535.
19. Kjeldskov, J. and Stage, J. New techniques for usability evaluation of mobile systems. *International Journal of Human-Computer Studies* 60, 5-6 (May 2004), 599-620.
20. Korn, M. and Zander, P.O. From Workshops to Walkshops: Evaluating Mobile Location-based Applications in Realistic Settings. In *Workshop on Observing the Mobile User Experience at NordiCHI 2010*, 29-32.
21. Lim, Y.-K., Stolterman, E. and Tenenberg, J. The Anatomy of Prototypes: Prototypes as Filters, Prototypes as Manifestations of Design Ideas. *ACM TOCHI* 15, 2, Article 7 (July 2008), 7:1-7:27.
22. McMillan, D., Morrison, A., Brown, O., Hall, M., and Chalmers, M. Further into the Wild: Running Worldwide Trials of Mobile Systems. In *Proc. Pervasive 2010*, Springer (2010), 210-227.
23. Oulasvirta, A., Kurvinen, E. and Kankainen, T. Understanding contexts by being there: case studies in bodystorming. *Personal and Ubiquitous Computing* 7, 2 (July 2003), 125-134.
24. Pascoe, J., Ryan, N. and Morse, D. Using while moving: HCI issues in fieldwork environments. *ACM TOCHI* 7, 3 (September 2000), 417-437.
25. Rogers, Y., Connelly, K., Tedesco, L., Hazlewood, W., Kurtz, A., Hall, R.E., Hursey, J., and Toscos, T. Why it's worth the hassle: the value of in-situ studies when designing Ubicomp. In *Proc. UbiComp 2007*, Springer (2007), 336-353.
26. de Sá, M. and Carriço, L. Lessons from early stages design of mobile applications. In *Proc. MobileHCI 2008*, ACM Press (2008), 127-136.