Like bees around the hive: a comparative study of a mobile augmented reality map

Morrison, Ann; Oulasvirta, Antti; Peltonen, Peter; Lemmela, Saija; Jacucci, Giulio; Reitmayr, Gerhard; Näsänen, Jaana; Juustila, Antti

Published in:
Proceedings of the 27th international conference on Human factors in computing systems

DOI (link to publication from Publisher):
10.1145/1518701.1518991

Publication date:
2009

Document Version
Early version, also known as pre-print

”Like Bees Around the Hive”: Interaction In the Wild with a Mobile AR Map

ABSTRACT

We introduce and present findings from field trials of MapLens, a mobile augmented reality (AR) digital-physical map system. In our trials we enlisted a mix of 37 early-adopters, environmental researchers, scouts and their families to use MapLens, to play an environmental awareness-raising location-based game. A comparative trial was run with a non-AR digital system. Analyses of videos, field notes, interviews, questionnaires and user-created content expose phenomena that arise uniquely when using AR maps in the wild. We report on how augmentation affects the way participants use their body and hands, manipulate the mobile device in tandem with the physical map, walk while using, and collaborate. We found that the MapLens solution facilitates place-making by its constant need for referencing to the physical, and in that it also allows for ease of bodily configurations for the group, encourages establishment of common ground, and thereby invites discussion, negotiation and public problem-solving. Its main potential lies not so much in use for navigation but in use as a collaborative tool.

Author Keywords

Augmented reality, mobile maps, mobile use, field study.

ACM Classification Keywords

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

INTRODUCTION:

Real-time processing of the mobile phone camera’s video stream has recently become computationally so efficient that it has enabled a host of augmented reality (AR) applications. In parallel, the topic has gained more attention in the HCI community. A central promise is that the information overlaid on the viewfinder supports understanding of one’s environment and its objects, as well as interaction with them. A unique characteristic of mobile AR is the *dual-presence* of information: aspects of the physical background (at which the camera is pointed) are represented simultaneously with extra information on the viewfinder.

Maps are one of the main application categories for mobile AR. The focus is in augmentation of *physical* maps with useful and interesting real-time information. Paper maps have a large static surface and AR can provide *a see-through lens* without forcing the user to watch map data *only* through the small “keyhole” of the display. Reported user studies have been conducted without exception in the laboratory (see Related Work). However, laboratory settings lack a number of dimensions that may or may not affect interaction. Particularly, in real world use the user is physically embedded in the environment to which the map and augmentation *refer* to. Moreover, the user may be walking and carrying out other tasks simultaneously and interaction may be carried out by not one but several people. The laboratory does not reflect the complexity of a real-use scenario.

Field studies of mobile AR maps are difficult to arrange due to the complexity of setting up the environments, and the problem of recreating realistic situations. Published work to date has utilised optical markers for tracking (e.g dotted maps). These change the appearance, hide information and require specifically printed maps. This is the first study that operates a markerless solution on a mobile phone. Our system, called MapLens, allows using a normal map that has not been visually altered.

To test the system, 37 users were recruited, 26 of which used MapLens (Figure 1) and 11 of which were assigned to a control group using DigiMap, a non-AR 2D map akin to Google Maps Mobile. The users operated in pairs or small teams in a pervasive game set in the center of Helsinki, Finland. Both systems allowed them to find information about the task targets as well as explore location-based media sent by other users. The game tasks forced the players to negotiate a range of different level tasks, carry multiple artefacts, and coordinate joint action, echoing real-world use. We col-

CHI 2009, April 4–9, 2009, Boston, MA, USA.

lected multiple kinds of data: video recordings and field notes, logs, interviews, and experience questionnaires.

We were surprised to find how MapLens invited the users to come together around the physical map and the mobile device, negotiating and establishing common-ground to solve the tasks together. By contrast, the non-AR 2D map was associated with problem-solving strategies that were more solitary and less collective. The data exposes the contributions of several technical characteristics, including stability, lighting, visualisation of targets, display size, type of and proximity to physical map, and so on.

Figure 1. MapLens in use with a paper map, overlaying digital information on screen. With the red square (centre) user locates and selects markers—as one user states—'*catches them'*.

RELATED WORK

The concept of *magic lens* was first introduced in 1993 [2] as a 'focus & context' technique for 2D visualisations and was later extended to 3D [20]. AR on handheld devices has been explored with different applications, including peephole applications where the background is used for positioning in virtual space [9, 24] and multi-user AR games [22]. A number of papers study target acquisition performance in magic lens pointing [4, 12, 17, 18].

Although carried out in the lab, there are two studies relevant to our work. Henrysson et al. [9] piloted positioning and orientation of 3D virtual objects using a mobile phone. They observed that users adopted a bi-manual strategy and sat down rather than stood up to stabilise the phone in hand. Reilly et al. [15] reported an exploratory study where 10 subjects performed pre-defined tasks on an RFID versus non-augmented PDA version. The effectiveness of the technique depended on the size of the map, information tied to it, and the needs of the user. The authors point out that the tasks required little or no spatial knowledge as the trial was conducted in a single location and involved no routes, landmarks, or navigation.

There are three critical aspects that are left to be addressed: interaction when embedded and mobile in the referred-to environment; interaction in pairs or teams; and suitability of mobile AR maps for real world use.

Pervasive games and locative media

We have chosen to use a location-based game as our evaluation method. Since early days researchers have pointed to the challenges of evaluating mobile AR in real settings. Limitations of the technological solutions have constrained researchers' ability to use real tasks and outdoor conditions. There is a growing interest in pervasive games as an evaluation methodology [10]. Recent work shows how pervasive games can be interwoven into daily life situations [1] and points out that results can bring forth aspects that are telling of issues beyond the game itself; such as interface design [14] or the user's learning [6]. We see no *a priori* reason for why mobile AR maps could not be evaluated similarly. The key challenge is to create a game that is not only motivating, but also engages the users with the environment in a way that can raise interesting phenomena that would perhaps not occur in task-based evaluation. Our game was designed to encourage players to be more aware of environmental issues while exploring their surroundings in a competitive but friendly game (for similar ideas, see [3]). The game involved several aspects of real-life situations, including multitasking, coordination of team effort, role-taking, sequential tasks, clear goals, feedback, social interaction [19], and time-urgency.

THE SYSTEM

MapLens is an application for Symbian OS S60 Nokia mobile phones with camera and GPS. When a markerless paper map is viewed through the phone camera, the system analyses and identifies the GPS coordinates of the map area visible on the phone screen. Based on these coordinates, location based media (photos and their metadata) is fetched from HyperMedia Database (HMDB). Markers to access the media by clicking the selected marker showing the thumbnail of the photo are then provided on top of the map image on the phone screen (Figure 1).

To help out selection in situations when there are multiple markers cluttered close together, a freeze function is provided: if there is more than one marker visible on the screen after the selection, the view is frozen with the markers being decluttered (pulled away from each other) so, that the user can more easily select the correct marker/thumbnail.

MapLens also functions as a camera for taking pictures that are uploaded in the background to HMDB. The user presses the * key to enter camera mode, 0 to capture a photo, and * again to return to MapLens. Photos are available for all within five minutes. By pressing 1 one is then able to see photos taken by other users. Pressing 1 again turns the feature off.

Markerless operation

MapLens uses predetermined map data files to identify the paper map and associate its visible area to geographical coordinates. Using this information, MapLens is able to position the media icons also on the edge of the paper map accurately. To accurately overlay information of the image of the map in the mobile phone's display, the 3D pose—

translation and rotation—of the phone's camera with respect to the map must be known. To track an image, we select distinct feature points in a representative template image and try to find these feature points again in the live image produced by the phone's camera. Because we do not modify the template image and do not require special fiducial markers to be applied, this is a so-called *natural feature tracking method*. Recent work in computer vision has given rise to a number of methods to accomplish this. However, our solution is among the first ones optimised to perform well on platforms with limited processing power [15].

The method implemented here [23] was optimised to operate on the N95 phone. In general the system operates at between 5 and 12 FPS, depending on the speed of motion of the camera allowing for interactive use. For this study a template image was used that allows operation from about 15 to 40 cm distance between the printed map and the camera. Tilt between the map and the camera is tolerated up to about 30 degrees, while in plane rotation is handled over the full range of rotations.

Figure 2. DigiMap version, Google Map with markers

A non-augmented system for comparison: DigiMap

As a comparison baseline for the user trial, we also instigated a *non-augmented map*, the design of which echoes Google Maps for mobile phones (Figure 2). No physical map was required, but we used the same map, red markers, and updated data to be switched on and off across both systems. We used standardised joystick phone navigation for *scrolling* across the map, using two buttons to control *zoom* in and out. However, our solution could not access the phone's camera, forcing users to switch from the web browser to the phone's native camera to take photos.

We developed both these systems, the technology and implementation for this project as a joint effort between XXXX, XXXX, XXXX and XXXX as part of XXXX.

THE FIELD TRIALS

Overview and timeline

In order to replicate a real-life scenario, we aimed to include real elements and tasks to imitate the kinds of circumstances that might usually be found around the use of this kind of technology. In order to achieve this, we designed a location-based game that required the users operate the systems and complete their tasks in a situated environment, that required they manage multiple levels simultaneously—with constant interruptions and shifts in focus, as well as conflicting distractions and divergent goals.

Three trials were held over three Sundays, in down town Helsinki, Summer, 2008. Prior, we piloted the game logic, timing, task difficulty, and interaction. Each trial was of an incrementally larger size. We had run a previous trial with an earlier prototype in Spring 2008. We included one team from this Spring trial in the first and third Summer trials to give comparative feedback on improvements. As well, five teams in the third trial tested DigiMap and the other five teams tested MapLens.

The participants

The first two Summer trials were comprised of largely professionals working in related fields, early-adopters, and researchers working with environmental issues. The third was comprised of scouts and their friends and families. Over these three trials we enlisted a total of 37 people with ages ranging from 7 years to 50 years, 20 females and 17 males. 21 had owned five or more mobile phones, with 22 owning or using regularly Nokia brand, and only one not familiar with or not owning a mobile phone. All phone owners used their phones for at least SMS and phone calls. Other self-reported information can be found in Table 1.

Table 1. Self-reported information from the participants.

	MapLens group 26	DigiMap group 11
Females + Males	19 + 7	1 + 10
Education	6 primary, 7 secondary, 13 tertiary	7 school, 3 secondary, 1 tertiary
ICT Knowledge	12 basic, 7 average, 7 expert	3 basic, 7 average, 1 expert
Hours of Technology Use week	6 > ten hours, 7 ten - 39 hours, 13 < 39 hours	4 > ten hours, 7 ten t-39 hours
Know Helsinki	8 no, 4 average, 14 yes	2 no, 2 average, 6 yes
Aware of Environmental Issues	9 average, 17 yes	3 no, 2 average, 6 yes
Navigation Skill	7 no, 19 yes	4 no, 7 yes
Used GPS	21 no, 5 yes	9 no, 2 yes

The game

The trials were run as location-based treasure hunt-type games. The game was designed to raise users awareness of their local environment. With the assistance of the technology the players followed clues and completed the given tasks within a 90 minute period, and in so-doing learnt about specific environmental concerns. We included three different prizes aimed at encouraging a variety of approaches to the game. One prize was for speed and accuracy—a more traditional approach to a game—another for the best photography, and one more for designing the best environmental task. An element of friendly competitiveness was established in the pre-phase game-orientation, and encouraged by promising prizes on return. Our intention was to focus and motivate our participants, as well as instigate time-pressure while they managed a broad range of multiple and divergent tasks simultaneously. Some tasks were more or less complex than others.

Table 2. Overview of the trial procedure

1. Pre-Phase. Fill in demographic & consent forms. Demo of technology and game. All familiarised themselves with devices.
2. Instructed Goal. "Complete as many tasks as possible in the allocated period." Awareness of other players assists users to navigate, compete with others for prizes.
3. The Game. The participants had four types of tasks to go through: *Inside the museum.* Task Type 1) Find clues and complete tasks. Task Type 2) Take photos of whole group. *Outside the museum.* Task Type 3) Find a recycle bin using software. Task Type 4) Locate and walk to sites and complete tasks such as water testing, sunlight photos. Record completion of tasks.
4. Post-Phase. Questionnaires and interviews.

The trial began at the Natural History Museum where players completed indoor tasks, two of which included follow-on components outside the museum (Table 2). We wanted the players to solve a variety of kinds of tasks (12 in all), some of which were sequential problem chains. For example, one museum task required information on an endangered Baltic seal; the follow-on task was to find the seals' home and calculate the carbon footprint by car, train and plane from an online site offering such comparisons. Provision for 20 minutes at an Internet café was included. How tasks were completed was up to players. As well there was no compulsory order for tasks to be completed in. Some tasks could be completed in several places, whereas others required visiting places in a certain order.

The game required players visit green areas in the city. One task was for the whole group to walk bare-foot in the grass, and upload a photo as evidence; another to gather a specific leaf (the leaf also found as a museum clue) and then take a sunlight photograph with a kit supplied, using water to develop the photo; another was to test a sample of sea water and a sample of pond water with a supplied kit for readings on Chlorine, alkalinity and pH balance. We added the 'taking-photo-of-whole group' component to many tasks to encourage physical proximity and team bonding.

Each team was handed a *kit* which contained seven objects in all (see Figure 3). By design, these objects required some coordination between team members to manage well. The large physical maps, expanding clue booklets, manipulating the phone over the map, writing in the clue book, the bag, meant that the participants needed to organise themselves into some kind of system of use. There were no ready-made solutions, in-situ creative problem-solving was required, and solutions varied according to the immediate environment—for example, a tree, a team mate or a near-by bench might be used as a steadying, leaning or resting prop.

Figure 3. Kitbags contained 7 items that needed to be managed: sunlight photographs, map, phone, water testing kits, voucher for internet use, clue booklet and pen.

Tasks were designed with a view to promote: internal and external group activities and awareness; negotiation of tasks and artifacts; 'noticing' and awareness of the environment; higher level task management; and awareness of physicality, proximity, embodiment and physical configurations around artifacts. There was particular emphasis on the mix of digital, and augmented with real and overtly tangible. These tasks were designed to facilitate proximate bodily configuration, to 'jolt' users away from small-screen absorption, and to remind the participants of their own corporeal selves [8].

For all trials, we ran thorough briefing sessions, to ensure all participants were familiar with game tasks and devices. When the teams left the briefing room each individual understood the immediate tasks and how to proceed.

Data collection

In the study we gathered data with a triangulation of quantitative and qualitative methods. Methods included collecting demographic data (Table 2) and ascertaining perceived experience with: technology, phones, use of maps, as well as knowledge of environmental issues and Helsinki center. Each team was accompanied through-out by one researcher taking notes, photographs and/or videos. On return from the game, participants completed a three-page questionnaire from Flow, Presence, and Intrinsic Motivation research to gauge reactions to the technology and the game [7, 19, 21]. This activity also focused participants on their experience in the trial, familiarising them with an extended vocabulary to better articulate those experiences. Each participant then described their experience, highlighting aspects that had caught their attention in semi-structured one-to-one recorded interviews.

Throughout the trial participants took photos as evidence of completing tasks. These images were synchronously uploaded from the phones, and assisted researchers to build an overview of activities undertaken during the trial.

OBSERVATIONS

Before moving on to the main subsections dealing with embodied interaction and collaboration, we briefly explain the general strategies of game play that we observed.

After the briefing session in the museum, the players headed for the clues—some even running—with many covering the same ground twice. Scout teams tended to ask museum guides or look for maps of the museum. Some teams split up while hunting, others stayed as a pack and were more systematic in their approach. Deciding a way to proceed, and more or less strategic game-plans unraveled in these early stages, varying between teams. Some teams, particularly those who knew each other well, divided the tasks with seemingly little effort or overt communication

Across the trials, we found that expert users' teams were more impartial in their turn-taking and role changing, whereas family members or friends tended to stay within their 'habitual' roles. For example, the youngest son in a family of four, automatically used the internet and where difficulties occurred was handed the phone, while the father and oldest daughter managed task order.

Photographing the environment

During the trials, the participants took a total of 184 photos. The DigiMap users were more eager to take photos as the average photos per team was 21.5 in comparison to 9.8 for MapLens teams. 36% of the photos taken by the DigiMap teams were task-related, with 45% by MapLens teams. Photos that do not count as task-related contained for example, photos of streets (7.6%), of parks (7.1%), of buildings (3.8%) and statues (3.2%). The DigiMap teams were somewhat more oriented to photograph the environment.

Embodied interaction

Comparing MapLens to DigiMap exposes several ways in which the systems both resource and constrain embodied interaction. By embodied interaction we refer here to the use of hands and body to position oneself, and the technology, in the context of other people and the environment.

We ask the reader to note that the figures presented from here on have been labeled with M when referring to MapLens and with D when referring to DigiMap.

Doing tasks with physical map versus the mobile map

In order to use MapLens, teams needed to use both the physical map and the device in tandem. For DigiMap teams, the use of the physical map was of course optional. Most MapLens teams used the physical-digital combination for identification of target location, but also for route planning (see Figure 4 left). As an exception to this a few groups unfamiliar with the surroundings used MapLens in two stages: first to identify the target destination and then the physical map alone to agree on the route to take (Figure 4 right). Three DigiMap teams did not use the paper map at all, or if they used it at the beginning, once it was put away in the bag, despite it having been useful, did not bring it out again. By contrast, MapLens teams were required to constantly negotiate this physical artifact to function in the game. They developed an expertise around handling the map, which in turn had a carry-on effect in the way they managed all the physical artifacts generally.

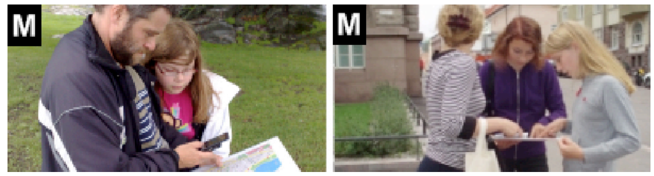

Figure 4. Most teams used MapLens (M) for both identifying the target and selecting the route. An exception is on right, a team that used the paper map having identified the target.

Holding the device

MapLens users typically held the device stretching out their arms because the camera needed to be held within the operating range of 15 to 40 cm away from the paper map. Moreover, the best light to view by was with sunlight on the map and the lens in shade. Importantly, by placing the device in this way, stretching one's arm, others could see what part of the map was being examined and at times contents on the display. We return to this issue, which we believe is central for encouraging collaboration on and around the map.

DigiMap users typically kept the device lower and closer to their body—a natural posture for holding a phone. However, this posture renders the phone more private (see Figure 5 right) and others cannot directly see the contents or reference points as with MapLens.

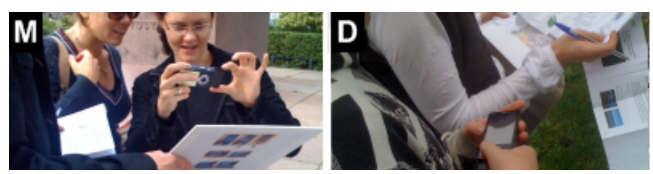

Figure 5. MapLens (M) was held in a way that it could be shared in the group, whereas DigiMap (D) users held the device more privately.

Use of two hands

The use of MapLens with the paper map often required two hands. The device was typically held in the dominant hand and the map in the other. Players also often used two hands to stabilise the phone, with another user holding the physical map (Figure 5 left). The players using MapLens had various items to carry with them and they often ended up gesturing with the device in the gesturing hand. While gesturing or organising their items, some players dropped the device on the ground (Figure 6 left). We observed this happening several times, but only with MapLens users.

The DigiMap players could use the device single-handedly, consequently they tended to have their non-dominant hand free, which allowed them to switch objects between their hands more flexibly (Figure 6 right).

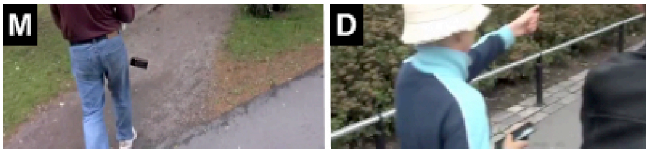

Figure 6. Use of hands with was different with MapLens (M). On left a MapLens user's drops his phone . By contrast, when using DigiMap (D), one hand is typically free.

Stabilising the map and lens

The players using MapLens had to stabilise the physical map and the device to be able to focus the lens properly. They often favored places where they were able to place the map on a table or bench. They also often laid the map on the ground or held the map for their group members (See Figure 7). This was a strategy to solve the problem of hand-tremble, which some MapLens users reported.

Figure 7. Stabilizing map surface for MapLens (left), then holding the device in two hands to mimimise tremble (right).

Turning and tilting the objects in hands

The paper map and the lens can be held in various orientations and alignments with the surrounding environment. When holding the paper map, MapLens typically kept the map aligned to north facing-up, and did not rotate the map around to align it with their orientation in the environment. As the map was somewhat cumbersome, rotating the map was more common when the map was supported by other players or surfaces, or when the map was on the ground.

The players using DigiMap occasionally turned the device—typically 90 degrees—for aligning the map with the environment. This may have been partly due to the smaller size of DigiMap setup that is easier to turn in hands. Another reason could be that the players struggled with establishing understanding of the map through the small screen size. Interestingly, about half of the players using MapLens kept the device horizontally, while the orientation of text and photos on the screen suggested vertical use.

Body posture

While the players using MapLens had to be relatively stable when using the system, DigiMap players were able to look at the map while moving around. Due to this we saw DigiMap users more often turning their body or glancing around while using the system (see Figure 8).

Figure 8. Turning to gaze the environment was more natural with DigiMap (D) that does not block view and constrain upper body movement as much as MapLens.

Walking while using

Seven of the eleven teams tried to use MapLens when walking, but all faced difficulties. In a typical scenario the team tried to use MapLens when walking but stopped doing so as the technology required steadiness to focus. There were two kinds of difficulties faced. First, even a very light trembling of the device makes MapLens difficult to use. Second, the participants' possibility to be aware of their immediate environment was challenged when using MapLens. One of our players was so engaged in looking at MapLens and the paper map that he walked into a lamp-post. These incidents indicate that MapLens does not support 'playing by moving,' but demands effort, forethought, and planning. Indicative of this, some teams used MapLens while waiting at traffic lights. In this way, they used their time well and it was possible to focus on using the technology without losing too much control over their immediate environment.

By contrast, difficulties of these kinds were not common in the DigiMap teams. Three of these teams used the system while walking, and one of the teams even ran while watching a map. However we did see a few cases of using MapLens while walking. A team of three young girls usually stopped to use 'MapLens & map', but as they began to run out of time, one of them walked more slowly behind the other two, who prevented her from running into anything (Figure 9 left). When she noticed something on the map, she called them to stop and look. As a group though, they did not use the technology when walking. Two other teams used MapLens while walking to watch the changing interplay as markers were picked up from the environment.

For MapLens players time spent walking was mainly used to get from one task to another, and to converse, or to discuss the last or the next task. Conversely for DigiMap teams walking was also an efficient time in the game, as it was inseparable from watching the map, and working out the next steps, so was less used for discussion.

Figure 9. Walking while using and bodily configurations. Left: Girls walk in front while one tries to read off MapLens (M). Center: MapLens (M) team negotiate where next. Right: One DigiMap (D) user reads the system while the other navigates.

Collaborative use

The previous section establishes salient differences in the manual and bodily operation of the two systems. We here turn to look at their implications on joint efforts. We start with analysis of handing over the phone as a physical object, then look at bodily configurations around MapLens, establishing common ground and place-making, and finally how conflicts were resolved.

Handing over phone

The handing over of the phone occurred more in the MapLens groups than in the DigiMap groups. As an example, in one instance with a MapLens expert-user group, we saw the one with the phone made an error about a place-name, and the next player while verbally corrected this error at the same time made a gesture of holding out her hand, and the phone was passed over. With a mother-son team where there was a constant struggle on which way to proceed. The boy retained DigiMap perhaps as a means to re-address the power imbalance. With a MapLens aunt and niece team, the only chance the niece had to use the 'MapLens & map' combination was when it was placed on the ground at the pool. She was the more competent user, but did not take it from her aunt, even though this meant they were less efficient in the game. The holder of the phone had the most agency in the team at that moment in time.

Bodily configuration

We observed teams negotiating together at all parts of the trial. The discussions did not only concern the task at hand and what the team should do next (and by which route) but also how to use the technology itself, as in Figure 9 (center), MapLens users in many instances gathered together around the physical map to use MapLens. The group members who did not have the phone gave instructions to the one holding MapLens on where to look. The need to hold the map stable did not give participants the freedom to do the navigation tasks while on the move (Figure 9 center), as it did for DigiMap where often one person was the "navigator" of the group searching things from the mobile, while others observed the environment and lead the way (Figure 9 right). This made DigiMap use more private and non-collaborative. Regardless participants using both systems found traffic lights a good place to negotiate where to go next and to find information from the maps or discuss clues.

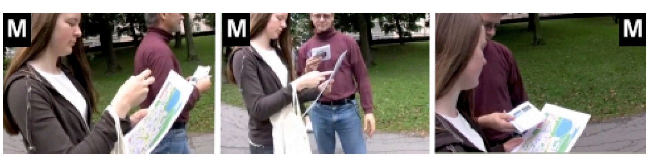

Figure 10. The physical map as a common ground, established by showing with the lens (M) and pointing with finger.

Establishing common ground

Given that the typical way of using MapLens involved a team gathered around the map and the main user gesturing on the map with the lens, establishing common ground was made easier for MapLens groups. By this term, we refer to shared understanding about the objects that are the focus of co-conversants' attention [5]. The location of MapLens on the paper map, and the contents that are revealed to others on its display, help others understand what the discussion is about without explicitly asking or negotiating. In Figure 10 a young woman browses the map by using MapLens. After finding an interesting place she suggests it to her father by pointing to it with her finger. The father proposes a nearby location instead and points to it by using the corner of a clue booklet.

The groups using DigiMap were not able to share the map that fluently. In Figure 11 a young boy is trying to identify a place by pointing to a relevant location on a screen and glancing around. After this he gestures towards the direction he suspects to be correct and hands the device over to his uncle, who then assesses the situation.

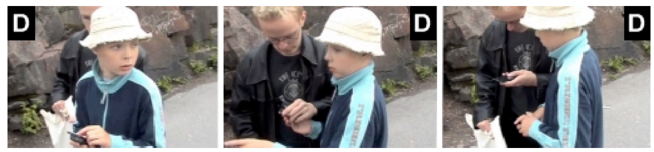

Figure 11. DigiMap (D) Attempting to share the map as common ground.

The physical paper map supported the players better in establishing a common understanding of the area and referring to different locations. Some players though found it challenging to identify the current location on the map with the focus of the lens, especially while it was being used by another player. The players using DigiMap often referred more directly by pointing at their surroundings.

The combination of the lens and the physical map provided the group a means to be collaborative in a more physical way. For example it was possible to pinpoint locations from the physical map either with a finger or a pen so that the participant using MapLens could easily target that point on the map (see Figure 12 left). As DigiMap use did not require using the physical map and the mobile phone screen is rather small in size, negotiations in DigiMap groups less often occurred with both trying to look at the mobile phone screen. Within a team of 2 close friends we observed constant pointing at the mobile screen, establishing common ground, others looked at the screen behind the "navigator's" shoulder (see Figure 12 right), but most often this was not done at all. Two DigiMap groups chose to use the physical map in addition to the digital map. For example, in one group a son searched for locations using DigiMap and either spoke aloud the options to his mother or pointed at them on the screen. The mother then used the physical map for a more detailed view of the surroundings.

Figure 12. Referring to objects by pinpointing. Left: Pointing with a pen while using MapLens (M). Right: pointing with finger from DigiMap (D) screen.

Place-making

The act of stopping walking, raising up the paper map and the lens, and gathering around for a while creates an ephemeral opportunity, isolated from the surroundings with the physical map and the bodies, to momentarily focus on a problem as a team. The phenomenon of place-making has been raised previously in the literature looking at mobile use of technology [11], and we encounter here a special multi-user form of it. Here, the physical map as a tangible artifact acts as a meeting point, a place where joint understandings can be more-readily reached by means of participants being able to see and manipulate and demonstrate and then agree upon action. The teams in pausing for discussion created a series of temporary spaces, places for collaboration. For example, they put bags down, swapped or rearranged objects they were carrying, and also stabilised the map and re-looked through MapLens to be sure they were on the right path. At this rapidly-made 'place' the tasks became again shared, negotiation and switching of roles often occurred and we witnessed a different kind of social usage in this temporary place. Other pedestrians walked around these 'places.'

Conversely the DigiMap teams only needed to stop at places that the tasks themselves dictated, the rest of the action and decisions and way-finding were mainly done while on the move.

Resolving conflicts

Even though the tasks required much group work, conflicts rarely occurred. Most problems were encountered while trying to locate things from maps. A failure to locate ended up usually in handing out the phone to some other group member who then tried to achieve the same task. This happened more often in MapLens groups as swapping the phone from one subject to another was easier as the augmentation usually happened while the group was standing still and close to each other. In DigiMap groups the 'navigator' less frequently handed out the phone to others, but it happened in some occasions, especially in groups of two composed of a parent and a child. These groups were also the ones that had most arguments on how the group should proceed in the game. For example, in one DigiMap group the mother did not allow her son to explore the environment, as she wanted their team to go as fast as possible.

QUESTIONNAIRES AND INTERVIEWS

Questionnaires

The participants filled in three questionnaires: a shortened version of MEC Spatial Presence Questionnaire (MEC-SPQ) [21], a GameFlow questionnaire based on [19] and an Intrinsic Motivation Inventory (IMI) questionnaire [7]. As Likert (ordinal) scale was used as a measure and Shapiro-Wilk's test revealed our data is not normally distributed, the Mann-Whitney U-test was selected to test differences between MapLens and DigiMap teams.

When comparing total Presence, Flow and Motivation score medians between MapLens and DigiMap participants, no significant differences were found. However, both groups scored above average on most items indicating that motivation, being present to the game and/or map system, and experiencing a sense of concentrated engagement was activated for users of both systems. When comparing individual Presence, Flow and Motivation items, significant differences were found. This may be due to questions addressing whether the system related to map system use, the game played or both (see Table 3).

Table 3. Questionnaire items having significant differences.

Item and Mann-Whitney U-test Significance: Presence 1-5 scale, Flow and Motivation 1-7 scale	System with higher median	System with lower median
*Items related only to **map system** use*		
Presence: I was able to imagine the environment and arrangement of the places presented using the map system well (*)	DigiMap MD=4.00	MapLens MD=3.76
Presence: It was as though my true location had shifted into the mapping system environment (*)	DigiMap MD=3.18	MapLens MD=2.29
Presence: I concentrated on whether there were any inconsistencies in this mapping system (*)	MapLens MD=5.00	DigiMap MD=4.00
*Items related to **both map system use and the game***		
Presence: The task and technology took all my attention (*)	MapLens MD=4.00	DigiMap MD=3.00
Presence: I felt I could be active in my surrounding environment (move, use the mobile phone and switch from task to task) (*)	DigiMap MD=5.00	MapLens MD=3.34
Flow: How to play the game and how to work the technology was easy (**)	DigiMap MD=6.00	MapLens MD=5.00
Flow: My skill level increased as I progressed (**)	DigiMap MD=7.00	MapLens MD=5.00
IMI: While I was working on the tasks I was thinking about how much I enjoyed it (*)	DigiMap MD=6.00	MapLens MD=5.48
IMI: I think I am pretty good at these tasks. (**)	DigiMap MD=6.00	MapLens MD=5.00
IMI: I found the tasks very interesting (*)	DigiMap MD=6.00	MapLens MD=5.00
*Items related only to the **game***		
Flow: The difficulty level got easier as the game progressed (**)	DigiMap MD=7.00	MapLens MD=4.31
Flow: I knew how I was progressing in the game as I was proceeding (*)	DigiMap MD=6.00	MapLens MD=5.35
Flow: I helped other players in other groups (**)	MapLens MD=2.08	DigiMap MD=1.00

Note: (*) = p<.05 and (**) = p<.01 .

As a general conclusion it can be stated that while the MapLens users felt confident using the technology and enjoyed the experience, the DigiMap users did so even more. The technology also enabled the DigiMap users to perceive their surroundings better than users of the MapLens system, who concentrated more on the technology as such, as well as being more focused on the game as a whole. Also MapLens users were socially active and more helpful of others. MapLens users were more focused and both groups scored high on sense of control, understanding requirements, interest and enjoyment.

Interviews: common participant descriptors

From the oral interviews we searched for recurrent descriptors (adjectives) in the participants' descriptions of their experiences. We found that MapLens users made 11 mentions of the word *stability* (compared to 0 with DigiMap). For example, "You need to be quite accurate; you need to be *stable* and you need to get the camera into the right position." Six MapLens users described the trial as *easy* compared to twenty-five instances of the word *easy* being used with DigiMap players. Here too, we find MapLens teams were more challenged by the technology: "At first it was difficult to find these dots. Maybe it was because we were not able to keep our hands stable enough. But after that we were able to find the dots, catch the red dots by using the square." The DigiMap technology was perceived as much *easier*, with zero problems with *stability*.

DISCUSSION

Mobile AR magic lenses do not work as singular objects, their operation is constrained in a particular way by their necessary material counterpart, for MapLens—the map. The central tenet to our findings is that seemingly minor details end up echoing down the sequential chain of events this aspect presents and essentially defines what mobile AR maps are good for. More precisely, 1) the stability of the feature tracking algorithm and therefore the stability required from the user, 2) the necessity of holding the map as the background surface, and 3) the operation being constrained within a proximity range of the paper map all influence both an individual's options in using the lens and the nature of collaboration in a team.

Embodied and mobile interaction. We found that the user needs to stretch out her arm to ensure MapLens is between 15 to 40 cm distance from the paper map for proper recognition of markers, and then position the 'map & MapLens' in relation to the environment, so that the map and lens are both in adequate lighting, but not direct sunlight. While the non-augmented digital counterpart of MapLens, DigiMap, is also susceptible to direct sunlight, it is much easier to cover such a small object with the palm of one's hand. Secondly, the use of MapLens, but not of DigiMap, effectively requires two hands, because either one has to steady the surface (the map) or use two hands to stabilise the phone in hand. For these reasons use while walking is not possible, whereas DigiMap was often used while on the go. We conclude that MapLens was more challenging to operate and it in fact restricted the mobility of an individual and the group. Moreover, the need for careful operation and focus on the 'surface & lens' restricted their attention to the surroundings. Users echo this description, describing interaction with MapLens as difficult and unstable.

Cooperation through place-making. However, MapLens use was not only *bad*. The typical team-level response was gathering around the map and the lens, a phenomenon we analysed as place-making: ***"like bees around the hive"***. Typically, one user held the map, another took over MapLens, and we see an establishment of bodily configurations in close proximity and negotiation of the next sequences of events occurring. The shared nature of using the device is in stark contrast to what we saw with DigiMap—typically one person taking the phone and leading the group with instructions and by showing the way. While MapLens users gathered around the map, we noted the importance of pointing to the physical map, with finger or pen and with MapLens itself. We argue that both support common ground. Keeping the lens on a particular area of the map that everyone can see reveals the holder's target of attention to others, and pointing at the surface (the map) makes this even more explicit. Keeping such 'bookmarks' on DigiMap was not that easy, because every scroll or pan would have necessitated updating the pointer's location as well.

User experience. As a general overview it becomes clear through the questionnaires, word mapping and photographic usage that MapLens users concentrated more on errors in the technology, but not the environment around them. Also MapLens users were more concentrated on the combination of the technology and the game—which involved problem-solving via negotiation, physical and social interaction. The way place-making affects attention to the task and technology, versus the surroundings is a plausabile explanation for this observation.

Toward real world applications

The underlying context of this work, and as the final theme we turn to the question of what mobile AR is good for and how to improve it. Laboratory studies are better suited for studying in detail the implications of tremble, map size, and visualization performance in close-loop interaction, but we want to raise the point that such improvements could also impact group use. From an individual user's perspective, robustness of the feature tracking algorithm is a worthwhile investment. However, in a cooperative setting it could lead to less swapping of the phone, and less need for the team to be involved in map-holding, which in turn would lead to less need for constant place-making activity, less interaction, discussion and negotiation. However, as one still needs to stretch out one's arm to hold out the phone and the map for correct working distance and visibility, then getting rid of tremble would have marginal impact on 'people coordination'. Also the implication with this technology is we can use any map, so if we also take away the need for the cardboard map, then suddenly we can use for example maps on billboards, maps in bus-stops. On horizontal surfaces one would still need to hold MapLens at the required distance from the map and ensure correct lighting for screen visibility, which in turn still invites pointing on common ground. Part of designing mobile AR maps is the design of the paper map and how it can be used in concert with the lens. In addition, we encourage designers to think about ways to support common ground even more. One way to do that would be to 1) increase display size and 2) make markers more salient for co-present others to see.

It would be trivialising to conclude that although MapLens' operation was manually more challenging it would have been "worse" than DigiMap, or that DigiMap's use was "less social" and therefore less desirable. Our conclusion is that while MapLens is not usable when walking, and MapLens users are less attentive of their surroundings when using it, cooperative group work benefits from the place-making that ensues and common ground that it supports. The wider implication for mobile AR research, then, is to look to establishing what kinds of tasks would require such mode of cooperation. These might include for example social gaming, public social tasks that require movement, interaction with the physical environment and information (maps or posters) and group puzzle solving scenarios—involving chains of complex sequential tasks—promoting discussion and focus.

REFERENCES

1. Bell, M., Chalmers, M., Barkhuus, L., Hall, M., Sherwood, S., Tennent, P., Brown, B., Rowland, D., Benford, S., Capra, M., Hampshire, A. Interweaving mobile games with everyday life. In *Proc. CHI 2006*, ACM Press (2006), 417-426.
2. Bier, E., Stone, M., Pier, K., Buxton, W., DeRose, T. Toolglass and Magic Lenses: The See-Through Interface. In *Proc. of ACM SIGGRAPH 1993*, ACM Press (1993), 73-80.
3. Benford, S, Flintham, M, Drozd, A, Anastasi, R, Rowland, D, Tandavanitj, N, Adams, M, Row-Farr, J Oldroyd, A, Sutton, J. Uncle Roy All Around You: Implicating the City in a Location-Based Performance, Proc. of ACE, 2004.
4. Cao, X., Jie Li, J., Balakrishnan, R. Peephole pointing: modeling acquisition of dynamically revealed targets. In *Proc. CHI 2008*, ACM Press (2008), 1699-1708.
5. Clark, H. *Using Language*. Cambridge University Press, 1996.
6. Costabile, M. F., De Angeli, A., Lanzilotti, R., Ardito, C., Buono, P., Pederson, T. Explore! Possibilities and Challenges of Mobile Learning. In *Proc. CHI 2008*, ACM Press (2008), 145-154.
7. Deci, E. L., Ryan, R. M. The "what" and "why" of goal pursuits: Human needs and the self-determination of behavior. *Psychological Inquiry 11* (2000), 227-268.
8. Heidegger, M. " "Poetry, Language, Thought" (Harper & Row, NewYork, 1971).
9. Henrysson, A., Ollila, M., Billinghurst, M. Mobile phone based AR scene assembly. In *Proc. MUM 2005*, ACM Press (2005), 95-102.
10. Jegers, K. Pervasive game flow: understanding player enjoyment in pervasive gaming. *Computers in Entertainment (CIE) 5*, 1 (2007).
11. Kristoffersen, S., Jungberg, F. L. "Making place" to make IT work: empirical explorations of HCI for mobile CSCW. In *Proc. International ACM SIGGROUP 1999*, ACM Press (1999), 276-285
12. Mehra, M., Werkhoven, P., Worring, M. Navigating on handheld displays: Dynamic versus static peephole navigation. ACM *Transactions on Computer-Human Interaction (TOCHI) 13*, 4 (2006), 448-457.
13. Mogi Mogi by Newt Games. Last accessed 09/09/2008. http://www.mogimogi.com/
14. Ohlenburg, J., Lindt, I., Pankoke-Babatz, U., Ghellal, S. A report on the crossmedia game epidemic menace. *Computers in Entertainment (CIE) 5*, 1 (2007).
15. Ozuysal, M., Fua, P., Lepetit, V. Fast Keypoint Recognition in Ten Lines of Code. In Proc. *IEEE CVPR 2007*, IEEE Press (2007), 1-8.
16. Reilly, D., Rodgers, M., Argue, R., Nunes, M., Inkpen, K., Marked-up maps: combining paper maps and electronic information resources. *Personal and Ubiquitous Computing 10*, 4 (2006), 215-226.
17. Rohs, M., Oulasvirta, A. Target acquisition with camera phones when used as magic lenses. In *Proc. CHI 2008*, ACM Press (2008), 1409-1418.
18. Rohs, M., Schöning, J., Raubal, M., Essl, G., Krüger, A. Map Navigation with Mobile Devices: Virtual versus Physical Movement with and without Visual Context. In *Proc. ICMI 2007*, ACM Press (2007), 146-153.
19. Sweetser, P., Wyeth, P. Gameflow: a model for evaluating player enjoyment in games. *ACM Computers in Entertainment 3*, 3 (2005), 1-24.
20. Viega, J., Conway, M. J., Williams, G., Pausch, R. 3D magic lenses. In *Proc. of UIST 1996*, ACM Press (1996), 51-58.
21. Vorderer, P., Wirth, W., Gouveia, F. R., Biocca, F., Saari, T., Jancke, F., et al. MEC spatial presence questionnaire (MEC-SPQ): Report to the European Community, Project Presence: MEC (IST-2001-37661), 2004.
22. Wagner, D., Pintaric, T., Ledermann, L., Schmalstieg, D. Towards Massively Multi-User Augmented Reality on Handheld Devices. In *Proc. Pervasive 2005*, Springer (2005), 208-219.
23. Wagner, D., Reitmayr, G., Mulloni, A., Drummond, T., Schmalstieg, D. Pose Tracking from Natural Features on Mobile Phones, In *Proc. ISMAR 2008*, IEEE Press (2008).
24. Wang, J., Zhai, S., Canny, J. Camera phone based motion sensing: Interaction techniques, applications and performance study. In *Proc. UIST 2006*, ACM Press (2006), 101-110.