Like Bees Around the Hive: A Comparative Study of a Mobile Augmented Reality Map

Ann Morrison[1], Antti Oulasvirta[1], Peter Peltonen[1], Saija Lemmelä[2], Giulio Jacucci[1], Gerhard Reitmayr[3], Jaana Näsänen[1], Antti Juustila[4]

[1]Helsinki Institute for Information Technology HIIT
Helsinki University of Technology TKK and University of Helsinki, Finland
[2]Nokia Research Center, Finland
[3]University of Cambridge, UK
[4]University of Oulu, Finland

ABSTRACT

We present findings from field trials of MapLens, a mobile augmented reality (AR) map using a magic lens over a paper map. Twenty-six participants used MapLens to play a location-based game in a city centre. Comparisons to a group of 11 users with a standard 2D mobile map uncover phenomena that arise uniquely when interacting with AR features in the wild. The main finding is that AR features facilitate place-making by creating a constant need for referencing to the physical, and in that it allows for ease of bodily configurations for the group, encourages establishment of common ground, and thereby invites discussion, negotiation and public problem-solving. The main potential of AR maps lies in their use as a collaborative tool.

Author Keywords

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

ACM Classification Keywords

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

INTRODUCTION:

Real-time processing of the mobile phone camera stream has become so efficient that it has enabled a host of augmented reality (AR) applications. A central promise is that information overlaid on the viewfinder supports understanding of one's environment and its objects. 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.

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

Figure 1. MapLens in use with a paper map, overlaying digital information on screen. The red square (centre) is used to select markers.

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. However, reported user studies have been conducted without exception in the laboratory (see Related Work). Laboratory settings lack a number of aspects that may affect interaction in real world use. Particularly, in real world use the user is physically embedded in the environment to which the map and augmentation *refer*. Moreover, the user may be involved in other tasks simultaneously and not one but several people may carry out interaction. Also traditionally, optical markers (e.g. dotted maps) have been used for tracking, which require specifically modified printed maps.

This is the first study that evaluates a markerless solution on a mobile phone out of the laboratory. Our system, called MapLens, allows using a normal, unaltered map. Thirty-seven participants were recruited for field trials, of which 26 used MapLens (Figure 1) and 11 used DigiMap, a digital 2D map akin to Google Maps Mobile (Figure 2). Pairs or small teams operated in a pervasive game set in the center of Helsinki, Finland. Both systems allowed they find information about the task targets as well as explore location-based media sent by others. The game tasks required players negotiate a range of different level tasks, carry multiple artefacts, and coordinate joint action, artefacts, and coordinate joint action, echoing everyday use. To understand interaction, we collected multiple kinds of data: video recordings and field notes, logs, interviews, and questionnaires.

We were surprised how MapLens invited participants to come together around the physical map and mobile device, negotiating and establishing common-ground to solve tasks. By contrast, DigiMap was associated with problem-solving strategies that were more solitary, less collective.

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 [32]. The NaviCam system [25] introduced magic lens on hand-held displays (see also [6, 17, 30]). Later on, mobile AR has been explored also with peephole interaction where the background surface is used for positioning the phone in virtual space [10].

Research in AR for paper maps has explored a wide range of output modalities. McGee et al. augmented a real map by automatically locating post-it notes placed on it to keep a computer model up-to-date [15]. Bobrich and Otto used head mounted displays in a video see-through AR setting to present 3D overlays of digital elevation models over a real map [5]. A projection-based system by Reitmayr et al. augmented a paper map directly with dynamic, geo-referenced information [24]. Transition to mobile devices has placed special demand on lightweight methods of localising. Reilly et. al. used RFID tags to associate locations on the map with digital information [26]. Rohs et al. describe computer vision-based method using sparse fiducial markers on a map [29].

User trials of any kind for mobile AR are scarce. Henrysson et al. piloted positioning and orientation of 3D virtual objects using a mobile phone [10]. They observed that the users sat down rather than stood up in order to stabilise the phone in hand. Reilly et al. reported a laboratory study where subjects performed pre-defined tasks on an RFID versus non-augmented PDA version [26]. Usability depended on the size of the map, information tied to it, and the task 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. Rohs et al. compared map navigation between joystick, static peephole and magic lens interaction [30]. The study showed switching of attention between the surface and background affects task performance, yet static peephole and magic lens clearly outperform joystick navigation.

Laboratory based studies have shed light to some aspects of mobile AR based interaction, but we find three critical aspects still need to be addressed: 1) interaction while embedded and mobile in the referred-to environment; 2) interaction in pairs or teams; and 3) suitability of mobile AR maps for real world use. Furthermore, typical laboratory experiments do not involve interruptions, involve very brief tasks, are completed by individuals, and do not involve physical aspects of the environment [30]. Conversely, the trials reported in this paper lasted 1.5 hours, involved a variety of inter-related and sequential tasks, and teams needed to interact with the physical environment as well as with other people in order to succeed in the game.

Pervasive games and locative media

There is a growing interest in pervasive games as an evaluation methodology [11]. 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 [19] or the users' learning [8]. We see no a priori reason why mobile AR maps could not be similarly evaluated. 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 (see [3, 13] for similar approaches). The game required managing multiple levels—with constant interruptions and shifts in focus—and involved several aspects of real-life situations including coordination of team effort, role-taking, sequential tasks, feedback, social interaction [31], and time-urgency.

THE SYSTEM

MapLens is an application for Symbian OS S60 Nokia mobile phones with camera and GPS. The map file used is a screen capture from Google Maps at 537x669 pixels, 72 dpi. 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 display (Figure 1).

To help with selecting when multiple markers are clustered close together, a freeze function is added. If more than one marker is visible on the screen after selection, then the view is frozen with the markers de-clustered (pulled away from each other) so the user can more easily select the correct marker/thumbnail.

MapLens also functions as a camera and photos are uploaded in the background to HMDB. The user presses * 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 can see photos taken by other users. Pressing 1 again turns that layer off.

Markerless operation

MapLens uses predetermined map data files to identify the paper map and associate its visible area to geographical coordinates. MapLens is then able to position media icons also on the edge of the paper map accurately. To overlay information on 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 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 led to a number of methods to accomplish this. However, our solution is among the first optimised to perform well on platforms with limited processing power [20].

Our implemented method [34] was optimised to operate on the N95 phone. 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 30 degrees, while in plane rotation is handled over the full range of rotations. We used the same Google map screen capture image for both the virtual and the physical maps, as well as for our DigiMap version. The physical map was printed onto an A3 size page, allowing white space around all sides of the map.

Figure 2. DigiMap version, Google Map with markers. Used by the control group in the trial.

2D digital system: DigiMap

As a comparison baseline in the final user trial, we added a digital version, the design of which echoes Google Maps for mobile phones (Figure 2). While no physical map was required in order to use this system, the same virtual map was used across both systems (the same physical map was supplied in the kit that MapLens users received), and the users had access to the exact same information—the red markers indicating extra information on clue places, and dynamically updating user data that could also be turned on and off via a layer in this system.

We used standardised joystick phone navigation for *scrolling* across the map, using two buttons to control *zoom* in and out. This solution did not access the phone's camera, so users switched from the web browser to the phone's native camera to take photos. We added the digital system to act as a control group in order to better understand results from earlier trials where we ran MapLens only.

Table 1. Self-reported background information

	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
Technology Use per 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
Environmental Awareness	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 FIELD TRIALS

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, with the final trial involving DigiMap. We had run a previous trial with an earlier prototype in Spring 2008. We included one team from this Spring trial in the first Summer trial and another in the third trial to give comparative feedback on improvements. As well, in the final third trial five teams tested the newly added DigiMap system to act as a comparative control group, while the other five teams tested MapLens. We wanted to see if there were differences in how people used two different systems for the same tasks.

The participants

The first two trials were comprised of largely professionals working in related fields, early-adopters, and researchers working with environmental issues. The third trial was comprised of scouts and their friends and families. The scouts teams were younger, predominantly male, less aware of environmental issues, with less expertise with technology, and understandably less tertiary qualifications (Table 1). Over the three trials, we enlisted 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 one unfamiliar with a mobile phone. All phone owners used their phones for at least SMS and phone calls.

In the third trial the scout groups were randomly allocated between the two systems, with a consequence that only one female was allocated to the group testing the DigiMap system. This introduced an imbalance that may impact the obtained results. We had actively sought a higher proportion of female users for the earlier trials to ensure gender differences in using technology were anticipated in early design and deployment stages [13, 16]. With younger and more male distribution in the third trial, we anticipated differences in patterns of use between same sex pairings, and participation styles [23], as well as language use [22]. Recent studies find a more collaborative approach in female pairings, differences in turn-taking and more or less aggressive styles between the genders [4, 16, 23]. Across all the trials, 19 users of MapLens were female with a higher education level, as well as more knowledge of ICT, navigation, local and environmental awareness and technology use, whereas DigiMap users were predominantly younger males (10) with less expertise in these areas. As well, seven males used MapLens and one female used DigiMap.

The game

The trials were run as location-based treasure hunt-type games designed to raise awareness of the local environment. With the assistance of the technology, the players followed clues and completed the given tasks within a 90-minute period. We included three different prizes aimed at encouraging a variety of approaches to the game: one for speed and accuracy—a more traditional approach to a game; another for the best photography; and another for designing the best environmental task. An element of friendly competitiveness was established in the pre-phase game-orientation, and encouraged with promising prizes. 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.

The game began at the Natural History Museum where players completed indoor tasks, two of which included follow-on components outside the museum. We wanted the players to solve a variety of types 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 outside the museum was included in order to achieve this. Another connected series was: find a leaf in the museum; find same leaf outside museum; take a sunlight photo of the leaf using water to develop (supplied in kit); test the pond water; test the sea water for chlorine, alkalinity and pH balance (supplied in kit); record all readings by uploading photos or entry into clue book and bring back results. 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. How tasks were completed and in what order was up to the players. Some tasks could be completed in several places, whereas series of tasks required visiting places in a certain order.

As GPS works only outside, participants found items in the museum by literally exploring the environment. The virtual maps for MapLens and DigiMap showed the same pictorial images as clues for the game; that is, an image of President Kallio statue showed on the map; a stage for more activity in the game. We also placed the same decoy images on both system maps; for example, air pollution toxicity levels, otherwise the game would be too easy. The photographs participants took outside were synchronously added to both DigiMap and MapLens versions of the map.

Each team was handed a *kit* that 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.

Game Design Rationale

Game 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 and tactile e.g. one task required team-lifting of a 27-kilo museum object. Such tasks encouraged physical proximity, team bonding and 'jolted' users away from small-screen absorption. We aimed to remind participants of their own phenomenological selves, interacting within the physical world [18], while synchronously accessing information via augmented or digital means. Tasks forced players to continually reorient their relationship to themselves as physical beings (and objects) within a world consisting of other physical beings and objects [18].

Data collection

In the study we gathered data with a triangulation of quantitative and qualitative methods. Each team was accompanied throughout 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 [9, 31, 33]. This 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

This section reports our observations on embodied interaction and collaborative use. Before moving on to the main observations, we briefly explain the general strategies of game play, users' photography, and game performance.

From here on in, we label figures and name groups with M when referring to MapLens and with D when referring to DigiMap. A limitation of this study is uneven gender distribution, where comparative interaction styles between the two technologies echo known gender differences with coordination and collaboration.

Game play

Player strategies

Overall, game strategies were similar between M and D teams. 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. 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 the scouts' family or friends tended to stay within their accustomed roles. For example, a younger son automatically used the Internet or was handed the phone when problems occurred, while father and daughter managed task order.

Table 3. Comparative numbers and types of photos

Photographs	All	DigiMap (D)	MapLens (M)
All Types	184	46.7%	53.3%
Average per team		21.5 photos	9.8 photos
Variation in teams		1 team 50 photos 1 team 20 3 teams 8	4 teams approx 10 1 team 1 photo
Task-related	76	36%	45.9%
Non-task-related	108	64%	54.1%

Photographing the environment

During the final comparative trial, participants took a total of 184 photos. We found differences in the kinds of photos taken and the average number of photos taken per team between our D and M users (Table 3). The types of photos were divided into a) task-related, meaning those captured in order to complete tasks and b) *non*-task related, meaning those captured from interest and unrelated to the game (e.g. streets 7.6 %, other parks 7.1%, or famous buildings 3.8% and statues 3.2%). We found D users took more non-task related photos and were more oriented to their surrounding environment, whereas M teams took more task-related photos and were more oriented to completing game tasks.

Game performance

Overall, we found M players took longer to complete the game, but were generally more fastidious with all tasks, and more accurate in reporting. Even though our results show D teams (scouts) were more aware of their surrounding environment, across the board they did less well in the task of designing a new environmental awareness task. As well, we found D non-map-related task completion details were not as accurately executed as with M teams. The D team that took 50 photos did not complete all game tasks, and two other D teams needed prompting to read thoroughly the clue booklet in order to attempt all tasks. These results support other findings that show M users were more embedded in the game itself. However, one scout M team (aunt-niece) only took one photograph, so did not successfully complete game tasks. Unsurprisingly, M early-adopter teams were the most fastidious and competent players.

Embodied interaction

Comparing M to D exposes several ways in which they 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.

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

Doing tasks with physical map versus the mobile map

In order to use M, teams needed to use both the physical map and the device in tandem. With D, the use of the physical map was optional. Most M 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 M 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). Two of the five D teams used the physical map for the entire game, with two others using this for much of the game, and one older team more experienced with mobile phones using the physical map in the training period only. M teams were required to constantly negotiate this physical artifact to function in the game, and developed an expertise around handling the map, which in turn had a carry-on effect in the way they generally managed all the physical artifacts.

Holding the device

M users typically held the device stretching out their arms because the camera needed to be held within the range of 15–40 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.

By contrast, D users typically kept the device lower and closer to the body—a natural posture for holding a phone. However, this posture renders the phone more private (see Figure 5 right) as others cannot directly see the contents or reference points as with M. Shading from the sun by use of one hand was possible with D, but this more private use also reflects that the team roles were less flexible, and for several teams revealed discomfort with close physical proximity.

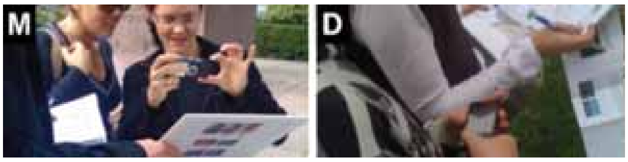

Figure 5. MapLens (M) was held in a way that it could be shared in the group, whereas DigiMap (D) users held the device more privately.

Use of two hands

The use of M 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, another the clue book etc (Figure 5 left).

All players had kit items to carry with them, and M players most often ended up gesturing with the device. While gesturing or organising their items, M users dropped the device on the ground (Figure 6 left) while D users most often dropped the clue book. In our March trial one user worked solo with M and completed the game within the allocated time, so one-handed use was proved possible.

After familiarising themselves with the system most of the predominantly younger male D players could use the device single-handedly, consequently towards the end of the game they tended to have their non-dominant hand free, which allowed them to switch objects between their hands more flexibly (Figure 6 right).

However, there were extenuating circumstances that may account for this. D users did not use the zoom in/out feature after their first experience of being *lost in the interface*, with one group handing it back to a researcher to fix and scrolling was an issue [30]. Most teams settled on using a zoomed out version where they could see most of the area they were active in, thereby avoiding joystick navigation.

Figure 6. Use of hands with was different with MapLens (M). On left a MapLens user's drops his phone . Conversely, when using DigiMap (D), one hand is free, but zoom not used.

Stabilising the map and lens

M users often had to stabilise the physical map and the device to be able to focus the lens properly. They 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 also in interviews (see also [30]).

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 MapLens+map combination can be held in various orientations and alignments to the surrounding environment. When holding the paper map, M users typically aligned the map to north facing-up, and did not rotate the map. Rotating the map was more common when the map was supported by other players or surfaces, or on the ground. Interestingly, about half of the players using M kept the device horizontally (as for a standard mobile phone), while the orientation of text and photos on the screen suggested vertical use.

D players occasionally turned the device—typically 90 degrees—for aligning the map with the environment. This may have been because the smaller size of D makes it easier to turn, or that D players struggled with reading the small screen size map.

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 (M).

Body posture

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

Walking while using

Seven of the eleven teams tried to use M when walking, but all faced difficulties of two kinds. First, even a very light trembling of the device makes M difficult to use. Second, the participants' possibility to be aware of their immediate environment was challenged when using M (e.g., a player walked into a lamp-post while looking at MapLens+map). Other variations were initiated. As a team of three young girls began to run out of time, one walked more slowly behind watching the device, with the others guiding her from running into anything (Figure 9 left). When she found something, she called them to look. Two other M teams persisted use while walking as they enjoyed the displayed screen interaction with markers picking up information from the environment. For M players time spent walking was used to get from one task to another, to converse, or to discuss the last or the next task. On the whole we found M does not support 'playing by moving,' but demands effort, forethought, and planning. Indicative of this, some teams used M while waiting at traffic lights.

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.

By contrast, difficulties with use while moving were not so common for D. Three teams used the system while walking, and one team of two young males even ran while watching the map. Therefore for D teams walking was also an efficient time to watch the map, and work out the next steps, so consequently was less used for discussion.

Collaborative use

We now look to joint efforts of users, starting with analysis of handing over the phone as a physical object. We then look at bodily configurations around M, practices of establishing common ground, and place-making.

Handing over phone

The handing over of the phone occurred more in the M groups than in the D groups. As an example, in one instance with expert users of M, we saw one user with the map including an error about a place-name. The next player verbally corrected this error and at the same time made a gesture of holding out her hand, and the phone was passed over. With a mother–son D team there was a constant struggle on which way to proceed. The boy retained D perhaps as a means to re-address the power imbalance. With an M aunt–niece team, the niece only got to use the MapLens+paper map combination 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 best use the technology, see Figure 9 (centre), M users in many instances gathered together around the physical map to use M. The group members who did not have the phone gave instructions to the one holding M on where to look. Needing to hold the map stable restricted movement (Figure 9 centre), unlike for D where often one person was the 'navigator' of the group searching things from the mobile, while others observed the environment and led the way (Figure 9 right). Bodily configuration around D use was separate and individual. The smaller screen and lower visibility meant less sharing occurred and dividing the roles took place earlier in the game.

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

Establishing common ground

Given that the typical way of using M involved a team gathered around gesturing on the physical map with the device, *establishing common ground* was easier for M groups. We noted a shared understanding around objects that are the focus of co-conversants' attention [7]. The location of M on the paper map, and the contents that are revealed to others on its display, help all to understand the discussion without explicitly needing to ask or negotiate. In Figure 10 a young woman browses the map by using M. After finding a place, she suggests it to her father by pointing to it with her finger. The father proposes a nearby location and points to it by using the corner of a clue booklet.

Figure 11. DigiMap (D) Users experienced difficulties while attempting to share the map as common ground.

D teams were not able to share the map that fluently. In Figure 11 a young boy tries to identify a place by pointing to a relevant location on screen and glancing around. After this he gestures towards the direction he suspects is correct and hands the device over to his uncle. This method hosts potential for more ambiguity and miscommunication.

The physical paper map supported the players better in establishing a common understanding of the area and referring to different locations. The combination of the Maplens+paper map provided a means to be collaborative in a more physical way with other objects: fingers, clue booklets, pens, and other components from the kit (see Figure 12 left). However, some M players found it challenging to identify the location on the map through the focus of the lens, especially while in use by another player. The D players often referred more directly by pointing at their 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.

For one D team we observed constant pointing at the mobile screen, establishing common ground. In another D team one looked at the screen behind the 'navigator's' shoulder (see Figure 12 right), yet in the three other D teams this did not ocurr. In one D group, a son searched for locations on the device 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. As such, the only female user in the D team used only the physical map due to poor legibility, difficulties with joystick navigation [30] and use [14, 23].

Place-making

Stopping, holding out the MapLens+paper map, gathering around for a short time created 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 with mobile use of technology has been raised previously in the literature [12], and we encounter here a special multi-user form of it. The physical map as a tangible artifact acts as a meeting point, a place where joint understandings can be more-readily reached and participants were able to see, manipulate, demonstrate and agree upon action. In pausing for discussion the teams created a series of temporary spaces, places for collaboration where they 'downed' bags, swapped or rearranged carried objects, stabilised the map and re-looked through M to ascertain progress. At this rapidly-made 'place', tasks were again shared, negotiation and switching of roles often occurred and we noted a different kind of social usage in this temporary place, with other pedestrians walking around these 'places.'

Conversely D 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 on the move or while stationary completing tasks.

Table 4. Questionnaire items showing significant differences between M and D groups

Item and Mann-Whitney U-test	**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: 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: 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 (**) = $p<.01$. Presence 1-5 scale, Flow and Motivation 1-7 scale.

QUESTIONNAIRES AND INTERVIEWS

Questionnaires

MEC-SPQ [33], GameFlow [31], and IMI [9] were used as basis for measuring user experience. 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 statistical differences between M and D teams.

When comparing total Presence, Flow and Motivation score medians between M and D participants we found 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 where questions addressed the system, the game played or both (see Table 4).

Three main conclusions can be drawn: 1) While M users felt confident using the technology and enjoyed the experience, the D users reported they did so even more. 2) D users were more aware of their surroundings than M users, who concentrated more on the technology, as well as being more focused on the game as a whole. 3) M users felt being more socially active and more helpful of others.

Interviews: Common participant descriptors

From the transcriptions of our interviews, we searched for recurrent adjectives in the participants' descriptions of their experiences. We found M users made 11 mentions of the word *stability* (and 0 with D). 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 M users described the trial as *easy* compared to 25 instances of *easy* being used by D players. Here too, we find M teams 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 catch the red dots by using the square."

DISCUSSION

The central tenet to our findings is that seemingly minor differences in embodied interaction imposed by the AR features echo down the sequential chain of events and essentially define how an individual user orients to her environment and how teams operate. We see 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 options in using the lens in relation to the environment, and the ways teams collaborate.

Our argumentation was based on comparison between MapLens and a standard 2D digital map. The typical team-level response for MapLens users was stopping movement and gathering around the MapLens+paper map combination, ***"like bees around the hive"***. Typically, one user held the map, another took over MapLens, and we saw establishment of bodily configurations in close proximity. We noted the importance of pointing to the physical map, with finger or pen and with MapLens itself and propose that both support establishment of common ground. As a general overview, it becomes clear through the questionnaires, word mapping, game results, and photographic usage that Maplens users concentrated more on the interface, 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 plausible explanation for this observation. Our conclusion is that although MapLens was more cumbersome to use for an individual, cooperative group work benefits from the place-making that MapLens ensues and common ground that it supports.

The contribution of this paper is in detailing interactions rather than just summing up which solution is better. The findings point out a couple of obvious opportunities for improving mobile AR interactivity in the wild. First, 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, getting rid of tremble would have marginal impact on coordination. Second, the implication with this technology is we can use any map, for example maps on billboards or in bus-stops. However, on horizontal surfaces one would still need to hold the device at the required distance from the map and ensure correct lighting for screen visibility, which in turn still invites pointing on common ground. As a conclusion, the collaborative support brought about by AR features would most likely not disappear by improving its technical functioning. Generally speaking, mobile AR features need to be designed and developed with a view to the 'real physical environment' they will be used within, not just the digital one. This means that field trials would become the standard for evaluation and experimenting, especially now that the technology is mature enough for outdoor conditions.

A broader implication for mobile AR research is to look to establishing what kinds of tasks would require the modes of cooperation that we have shown MapLens to support. 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.

ACKNOWLEDGEMENSTS

We thank our reviewers for their generous input. MapLens and DigiMap were developed by University of Oulu, Graz University of Technology, University of Cambridge, and HIIT, as part of the EU FP 6 funded project IPCity. We acknowledge Academy of Finland and Australasian CRC for Interaction Design (ACID) for support.

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. *Proc. CHI '06*, ACM Press (2006), 417-426.
2. Bier, E., Stone, M., Pier, K., Buxton, W., DeRose, T., Toolglass and Magic Lenses: The See-Through Interface. *Proc. SIGGRAPH '93*, 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, In *Proc. ACE*, ACM Press (2004).
4. Benford, S., Bederson, B., Åkesson, K., Bayon, V., Druin, A., Hansson, P., Hourcade, J.P., Ingram, R., Neale, H., O'Malley, C., Simsarian, K.T., Stanton, D., Sundblad, Y. and Taxén, G. Designing storytelling tech-

nologies to encourage collaboration between young children. *Proc CHI 2000* ACM Press (2000), 556-563

5. Bobrich, J., Otto., S. Augmented maps. In *Geospatial Theory, Processing and Applications*, *Proc. ISPRS Commission IV*, WG7, (2002).
6. Cao, X., Jie Li, J., Balakrishnan, R. Peephole pointing: modeling acquisition of dynamically revealed targets. *Proc. CHI 2008*, ACM Press (2008), 1699-1708.
7. Clark, H. *Using Language*. Cambridge University Press, 1996.
8. Costabile, M. F., De Angeli, A., Lanzilotti, R., Ardito, C., Buono, P., Pederson, T. Explore! Possibilities and Challenges of Mobile Learning. *Proc. CHI 2008*, ACM Press (2008), 145-154.
9. 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.
10. Henrysson, A., Ollila, M., Billinghurst, M. Mobile phone based AR scene assembly. *Proc. MUM 2005*, ACM Press (2005), 95-102.
11. Jegers, K. Pervasive game flow: understanding player enjoyment in pervasive gaming. *Proc CIE 5*, 1 ACM Press (2007), Article 9.
12. 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
13. Konopatzky, P., Scholz, M., Reißig, S., Verhoeven, P., Löchtefeld, M., Schöning, J., Krüger, A. xChase: A location-based multi-user pervasive game using a lightweight tracking framework. *Proc. FnG 2008*. Springer, (2008).
14. Liff, S., Shepherd, A. An evolving gender digital divide? Oxford Internet Institute, *Internet Issue* 2, (2004).
15. McGee, D. R., Cohen, P. R., Wu, L. Something from nothing: Augmenting a paperbased work practice via multimodal interaction. *Proc. DARE 2000*, ACM Press (2000), 71–80.
16. Maurin, H., Sonnenwald, D. H., Cairns, B., Manning, J. E., Freid, E. B., and Fuchs, H. 2006. Exploring gender differences in perceptions of 3D telepresence collaboration technology: an example from emergency medical care. *Proc. NordiCHI* ‘06 ACM Press (2006), 381-384.
17. Mehra, M., Werkhoven, P., Worring, M. Navigating on handheld displays: Dynamic versus static peephole navigation. *Proc. TOCHI* ACM Press (2006), 448-457.
18. Merleau-Ponty, M., The Visible and The Invisible, Lingis, A (Translator) Northwestern University Press; 1 edition, 1969.
19. Ohlenburg, J., Lindt, I., Pankoke-Babatz, U., Ghellal, S. A report on the crossmedia game epidemic menace. In *Proc. CIE 5*, 1 (2007).
20. Ozuysal, M., Fua, P., Lepetit, V. Fast Keypoint Recognition in Ten Lines of Code. *Proc*. *IEEE CVPR 2007*, IEEE Press (2007), 1-8.
21. Pasman, W., Woodward C., Hakkarainen, M., Honkamaa, P., Hyväkkä, P., Augmented reality with large 3D models on a PDA: implementation, performance and use experiences. *Proc*. *SIGGRAPH* 2004. ACM Press, 344 – 351.
22. Poynton, C. Language and gender: Making the difference. New York: Oxford University Press, 1989.
23. Prinsen FR., Volmann MLL., Terwel J, 2007, Gender-related differences in computer-mediated communication and computer-supported collaborative learning. *JCAL,* Wiley (2007), 393-409.
24. Reitmayr, G., Eade, E., Drummond, T. W. Localisation and interaction for augmented maps. *Proc. ISMAR 2005*, IEEE Press (2005), 120–129.
25. Rekimoto, J. Navicam: A magnifying glass approach to augmented reality. *PRESENCE - Teleoperators and Virtual Environments*, 6(4). MIT Press (1997), 399–412.
26. Reilly, D., Rodgers, M., Argue, R., Nunes, M., Inkpen, K., Marked-up maps: combining paper maps and electronic information resources. *Proc. PUC 10*, 4 (2006), 215-226.
27. Rogers, Everett M. New Product Adoption and Diffusion. *JCR*, 2 (1976), 290 -301.
28. Rohs, M., Oulasvirta, A. Target acquisition with camera phones when used as magic lenses. *Proc. CHI ‘08*, ACM Press (2008), 1409-1418.
29. Rohs, M., Schoening, J., Krueger, A., and Hecht, B. Towards real-time markerless tracking of magic lenses on paper maps. *Proc. PERVASIVE 2007*, 69–72, ACM Press (2007).
30. Rohs, M., Schöning, J., Raubal, M., Essl, G., Krüger, A. Map Navigation with Mobile Devices: Virtual versus Physical Movement with and without Visual Context. *Proc. ICMI 2007*, ACM Press (2007), 146-153.
31. Sweetser, P., Wyeth, P. Gameflow: a model for evaluating player enjoyment in games. *Proc. CIE 5,* ACM Press (2005), 1-24.
32. Viega, J., Conway, M. J., Williams, G., Pausch, R. 3D magic lenses. *Proc. UIST ‘96*, ACM Press (1996), 51-58.
33. Vorderer, P., Wirth, W., Gouveia, F. R., Biocca, F., Saari, T., Jancke, F., et al. *MEC-SPQ*: *Report to the European Community*, Project Presence: MEC (IST-2001-37661), 2004.
34. Wagner, D., Reitmayr, G., Mulloni, A., Drummond, T., Schmalstieg, D. Pose Tracking from Natural Features on Mobile Phones, *Proc. ISMAR 2008*, IEEE Press (2008).