iPhone *in vivo:* video analysis of mobile device use

Barry Brown
Mobile Life @ Stockholm University
SE-164, Kista, Sweden
barry@mobilelifecentre.org

Moria McGregor
Centre HCI Design
City University, London
moira.mcgregor.1@city.ac.uk

Eric Laurier
School of Geosciences
University of Edinburgh
eric.laurier@ed.ac.uk

ABSTRACT

Despite the widespread use of mobile devices, details of mobile technology use 'in the wild' have proven difficult to collect. This paper uses video data to gain new insight into the use of mobile computing devices. Our new method combines screen-capture of iPhone use with video recordings from wearable cameras. We use this data to analyse how mobile device use is threaded into other co-present activities, focusing on the use of maps and internet searches. Close analysis reveals novel aspects of gestures on touch screens, how they serve 'double duty" - both as interface gestures but as as resources for ongoing joint action. We go on to describe how users 'walk the blue dot' to orientate themselves, and how searches are occasioned by the local environment. In conclusion, we argue that mobile devices - rather than pushing us away from the world around us - are instead just another thread in the complex tapestry of everyday interaction.

Author Keywords

Video methods, smartphone use, mobility, ethnography.

ACM Classification Keywords

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

General Terms

Human Factors

INTRODUCTION

Any observer of technology cannot have missed the ongoing growth in mobile device use. While conventional keyboard and mouse based computing has hardly disappeared, the mobile phone's increase in functionality, sophistication and raw sales, has changed fundamentally the assumed platform for technology use. By some estimates mobile internet-use will supersede fixed internet-use by 2015 [17]. Computing use appears to be bifurcating into the existing mouse and keyboard based interaction (usually on a laptop) contrasting with touch interaction on smaller hand-held devices such as tablets and mobile phones.

This shift has contributed to the growing interest in mobile computing within HCI. Yet we would argue that we still have only schematic understandings about how mobile devices are actually used. In part this is because of the methodological challenges of working with devices that are by their very nature lightweight and used in a range of uncontrollable, distributed settings.

To address this, the paper here presents a study of mobile device use drawing on video recorded from multiple portable wearable cameras. We combine this video with recordings from their smartphone made with screen capture software. This allows us to relatively unobtrusively record device use and surrounding user activity, in naturalistic settings without the presence of potentially distracting equipment. We studied 10 city-trips, both visitors and locals on an ordinary non-work day. Participants made use of their iPhone as they would normally, except recorded on video. Through this we collected just under 24 hours of video of mobile activity, and around four hours of video of mobile device use, with recordings of the surrounding environment, conversations and interactions with others.

Existing research on mobile device use emphasises the interaction with the device, the importance of particular applications and gross patterns in application use [1, 3, 8, 18]. In contrast, the study here focuses on the relationships between the social and physical environment, and how device use responds and reshapes action in that environment. The use of the mobile devices captured was shaped by the contingencies of particular situations - such as checking a map while ordering a drink, having a conversation around a web search, or visibly and accountably dis-attending to others while reading content on a mobile device. This data would have been difficult to collect using existing mobile research methods.

We focus on two popular uses of mobile devices: *information search* and *map use*. We describe the occasioned nature of search on mobile devices, where information searches happen just before or at the point when that information is needed to be used, drawing on the environment and conversation. With our examination of map use we describe how directions come to be oriented to the local environment and used to navigate and find suitable destinations. The challenges of users in overcoming the inaccuracies of sensors and maps are also outlined.

This new method, while relying on earlier analytic work using video and interaction analysis, provides a new viewpoint on mobility and mobile device use. This allows us to document device use as not simply something done while co-present with others but action intimately connected with and dependent upon others in the environment. We conclude the paper by arguing for seeing

CHI'13, April 27 – May 2, 2013, Paris, France.

mobility not in terms of physical movement or device design, but in terms of the embedding of device use in the world of co-present others.

PREVIOUS WORK

The massive growth in mobile phone use, and the more recent explosion in mobile device functionality, has been one of the most notable changes to our modern world. As Merchant puts it "Mobile phones have rapidly been absorbed into the fabric of our day-to-day lives", [22]. A range of statistics can be drawn upon to argue for the importance of mobile phones as a computing device. Smartphones outsell PCs, and recent surveys have reported that mobile phones are used for longer than all other computers combined [17]. Tablet use is also one of the quickest growing, new, multi-purpose technology forms, followed by the growth in ebook readers and other devices with a similar form factor. Indeed with laptop computers outselling desktops, one might even say that the majority of computing is already 'mobile'.

While smartphones can be relatively simple to demarcate, the question of what devices are 'mobile', and what mobility means for research can be harder to define. One approach is to see mobility in terms of the mobility of users themselves. For Brown and O'Hara [6], 'mobility' is a concept that has analytic value as a worker's concern. This research documents how 'mobility' is something that has to be dealt with in terms of managing place - when and where work will be done, who will be there, and what artefacts are need to be deployed to make places 'work-place-able'. In a similar way, Luff and Heath [20] discuss mobility yet focus instead on the manipulation of artefacts in collaborative settings. Here, their interest is on what they term 'micro mobility': the ways in which information - particularly paper documents - come to be deployed in face to face interactional settings. The dynamic is between the co-present revelation and obscuring of information.

Alternatively, the predominant user interface paradigm built into different devices can be seen as providing a definition of mobility. Perhaps the most notable recent change has been the growth in finger-based interaction with smaller mobile devices. In terms of user interaction, devices can be divided between 'touch' systems (predominately iOS and Android) and keyboard and pointer based systems (predominately Windows and OSX). With a market in flux this distinction is open to revision - although it does highlight a fairly distinct class of touchscreen devices with distinct forms of use. As we will argue, this demarcation of mobile devices as touch systems also reveals that we have relatively little understanding of the differences in how these touch systems are used.

A number of classic studies have documented the use of mobile phones, however this work mainly discusses the use of conventional mobile phone features, such as text messages or phone calls. Weilenmann discusses shared phone use [32], Murtagh describes mobile phone use in public [24], and Taylor [2] documents some of the broader interactions around mobile phone use amongst teenagers being some of the most notable. Palen's voicemail-based diary-method examined issues surrounding cost and plan usage, relatively neglected for research since then [26].

Focusing specifically on smartphone use, three studies are revealing. The first is Barkhuus and Polichar's study of smartphone use [1], where they document the importance of smartphones as combining multiple functionality in a relatively 'messy' way. While this work documents revealing aspects of use, its reliance on interviews and diaries means that its description of use is indirect. There is a lack of description of actual cases of smartphone use.

In contrast, Böhmer et al. [3] used logging of application launches on the Android platform to present a detailed view on which applications are used on smartphones, alongside an analysis of time and length of use. An interesting concept introduced in this work is 'application chains' - where users switch from one application to another in a 'chain' of activity. While they report that the majority of sessions consist of only one application launch, 32% of sessions have two or more applications. Although this paper gives us an important view on what applications are being launched, there is relatively little information on what is then being done with the applications themselves.

Techniques of device logging such as this have much in common with experience sampling, such as MyExperience and Momento [7, 9]. While these systems pioneered using instrumentation and user alerts on mobile devices to collect data about use and broader user interaction, they have not been used as part of studying modern mobile device use. One exception is the work of Church and her collaborators [8], which has focused on mobile search. This research makes extensive use of experience sampling (prompted through SMS messages), and web-based surveys of mobile search. This work is particularly strong in understanding the context and motivation behind mobile search. For example, reporting on the importance of conversations as a prompt for web search. Yet it also suffers from analysing reported, rather than actual, use.

Previous research thus leaves much about modern mobile device use unanswered. In particular, there is little in the way of understanding the details of usage of particular applications and what shapes usage beyond simply what applications are used when. For example, what prompts the use of particular applications at particular times and locations? What purposes are different applications put to? How might application use be influenced by the setting, what role does task have in use, and how might smartphone use influence action more broadly? To be answered, these sorts of questions demand a different method.

USING VIDEO TO STUDY MOBILE DEVICE USE

As we can see, the predominant methods in studying mobile device use have been either *post-hoc* interviews, diary, logging or experience sampling based. While these methods have considerable strengths they face two immediate challenges. The first is a lack of of coverage. Interviews take place after use, diary reports are made just after use, and experience sampling depends upon sampling situations of use. These reports are not likely to capture every situation of use. Respondents prioritise certain incidents over others, or simply forget brief uses [29]. This issue is compounded by the lack of detail. Again, post-hoc interviews offer only outlines of usage, filtered again through what an interviewee may consider relevant or important. While diaries can be used to ask detailed questions, there are limits to what can be practically asked of an informant [25]. It is not that these methods are not useful - of course, any method has its shortcomings and limitations - but that these absences limit our view on device use since much of the data is recorded *after* the point of use. In contrast, while log or instrumentation studies collect data at the point of use, these logs offer only a schematic outline of device usage. This is compounded by analysis which seems to describe general use, rather than attend to understanding specific cases of use.

Our goal for this paper was therefore to experiment with new methods that might capture more of the situation of usage, and allow us to take a different analytic frame. Video recordings have been one longstanding method used in CHI and CSCW for a range of different analytic purposes, but particularly to document interaction with and around technical artefacts [13, 16, 31]. This research has extensively studied interaction around screens, capturing on-screen interactions as part of ongoing complexes of work practice (e.g. [14]). Video has proven valuable in illuminating aspects of activities neglected by previous methods, particularly drawing on conversation analysis to understand technology in use. In some senses video forces attention on the moment-by-moment production of technologically mediated action. A recent example is our recent research on driving with GPS [5] where cameras mounted inside cars were used to record drivers navigation. This work elaborated on the detailed 'problems' overcome by drivers using GPS, problems on the whole missed by more experimental studies.

Using video methods to study mobile use, however, presents considerable challenges. Since the users themselves are largely mobile, fixed cameras are of limited use. Small digital wearable cameras, while potentially portable, also have a number of limitations in battery life and flexibility. Camera angles can be vital in capturing the situation of use and mobile cameras are more difficult to aim, constrain or control. While fixed cameras can be pointed at desktop computer screens this is not as easy with mobile devices which may be held at any angle, in sunlight, and may have particular software limitations.

More broadly, 'sessions' of desktop computer use can be organised as part of an experimental task. Yet mobile device use is something that takes place sporadically throughout the day and is frequently dependent on the local setting. Constraining device use to a specific place or time would likely distort somewhat the nature of use. Applications which rely upon the environment or relevant setting are much harder to study as part of a pre-arranged experiment.

METHOD

To address some of these issues we developed a relatively lightweight video method that made use of wearable cameras, combined with screen recording software. This method allows us to gain something of a new perspective on mobile device use and, importantly, lets us study in detail how the environment and device are connected in use. Our participants wore multiple portable wearable cameras that recorded their actions relatively unobtrusively (figure 1). These videos were combined with phone screen recording software to allow us to unobtrusively capture the use of smartphones. We combined this with experimenting with studying multiple participants together, each wearing portable cameras, and obtaining multiple viewpoints of device use, with video recordings of the user of a device (from their companion) alongside video recorded by the device user, and a recording of screen use. For individual users, one of the authors accompanied them for the first hour or so of their visit. This allowed us to collect a range of videos of both collaborative and solo device use.

Figure 1: Participant is wearing camera in map bag around her neck. The left picture is taken from one wearable camera, with the image from the iPhone recording and the second camera on the right.

In our previous work we have long struggled with obtaining good video images of smartphone screens. Due to glare, video recording smartphone screens outdoors is very difficult. We dealt with this problem by developing software that unobtrusively records the mobile display in real time without interfering with its use. The software takes a static image of the screen four times per second, and then uploads the images to a server. The frames are then combined to produced videos of device use. While this limits the exact timing of some recorded actions, it was more than sufficient for our analytic work. Combining this screen capture with the wearable cameras we aimed to produce real time recordings that would give us access to the user actions and the environment on and off-screen.

We recorded a diverse set of everyday practices of smartphone use, centred around the use of smartphones as part of a city 'day trip'. We recruited a mix of locals and visitors to Stockholm, Sweden who regularly used an iPhone branded smartphone. In total fifteen mobile device users participated in the study, five recruited as individuals and five recruited as couples, with ten trips studied in total.

The participating smartphone users were recruited from advertisements in local cafes, and adverts on visitor websites. We recruited a mix of five groups of locals and five groups of visitors, volunteering to let us record them during a 'typical' day out. Four of our individuals were female, and one male, and two of our couples were male/ female, and three all female. Our participants were recruited with the request they take a 'typical day out in Stockholm'. All the participants were fluent in English, and we asked them to interact in English - although some participants switched to other languages for parts of their trip.

Our screen capture app was written for the iPhone. Due to a range of issues, seven of the ten participants made use of an iPhone supplied by ourselves. When running the trial, two participants carried Android phones which could not have their screens recorded. Another three participants requested that we did not install any software on their phone and two participants did not have data access enabled on their phone. For these reasons, only three of our participants had our recording software installed on their own phone. For the remaining seven participants we instead supplied them with our own iPhone with the recording software pre-installed. For the participants that went out carrying our device, they nearly exclusively used our device which was recorded. In one case a group of two participants used both devices - one using their own phone, and one our iPhone. In two additional cases text messages and phone calls were conducted on the users' own device, but the supplied recording device was used for their other smartphone use. To record their interactions we equipped participants with small GoPro portable cameras worn over their neck in map-bags. Audio was recorded with a small external microphone attached to the outside of the bag. We combined the screen recordings, with the multiple camera angles to produce composite videos of each participant's trip. These ranged in length, but were limited by the battery life of the GoPro cameras (around four hours).

The analysis of the data grows out of ethnomethodology and conversations analysis. Our goal was not to provide a statistical breakdown of the uses of the mobile device. In many ways this is already well catered for in the existing log-based literature. Rather our goal was to better understand the details of particular situations of use. In individual and group data analysis sessions we surveyed the journeys, editing the 24 hours of video down to 205 one to two minute clips for focused analysis, of which 12 were transcribed and analysed in depth. Our goal in this analysis was to understand both in a broad sense how mobile device use was organised, but also how specific incidents of use unfolded, and what contingencies users brought to bear.

RESULTS

While obviously the device usage we recorded was fashioned to an extent by our recruitment, our participants conducted a range of different visits (figure 2). In doing so they also used a wide range of applications including: Yelp, Facebook, Instagram, Google Maps, Phone, Web browsing and SMS. Users provided with our own phone did not use the iPhone's Facebook or mail clients, but accessed these services via the websites instead. Prominent in the smartphone usage captured was the use of the internet to browse for information, and the iPhone's maps application to find and navigate on foot and public transport.

1. Visit to ferry and music concert
2. Walk through the city to collect bike from repair shop, followed by cycling to music concert.
3. Lunch followed by shopping trip
4. Visit to botanical gardens
5. Visit to church and walk across the Stockholm
6. Visit to local town outside city
7. Visit to Reimersholm and then to the city hall
8. Visit to Old Town to watch changing of the guard
9. Visit to food hall
10. Shopping trip and visit to gardens

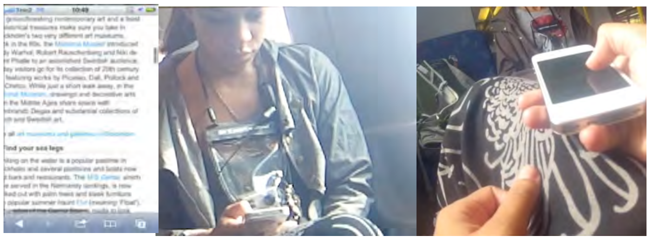
Figure 2: list of activities

Multiple involvements: Divergent and convergent

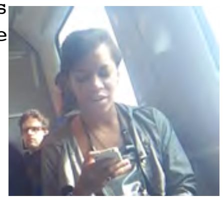
One of the most prevalent features of the video recordings is the extent to which device use takes place alongside involvements with co-present activities and other people. These interaction can be characterised as 'multi-modal', in that interaction takes place with the device but also through spoken conversation, or 'talk-in-interaction', [27]. What is also clear is that the device use itself shapes the interactions that unfold. If mobile device use takes place with others who are co-present, device use is frequently brought into the interaction - what Mondada [23] characterised as 'convergent' interaction. In contrast we can identify 'divergent' interactions [23] where the use of the device is kept relatively separate from interactions with others.

In figure 3, one tourist is talking to a researcher while on the subway. She pauses in conversation before moving on to interacting with her phone to read some details about the section of Stockholm she is visiting. She returns to the conversation by reading out a name from the website she has found, and then returns to reading. In this clip the reading of the webpage takes place mainly without interaction with the researcher who is sitting in the seat opposite. Moreover, the visitor 'reading' is something that is visible in their orientation to the device. In the short exchange with their companion, the user lifts their head and reads aloud, keeping her gaze on the phone. In this clip the involvement of the participant with the conversation and reading are kept separate - her attention, gaze, and body compartment used to keep apart her engagement with the device from her participation with her companion.

In studies of talk and interaction one recurrent concept is that of 'multi-activity' settings where participants are engaged in multiple streams of activity simultaneously [11, 15]. One well researched example of this is in driving and conversation - where the activity of driving plays a role in regulating the flow of conversation. Drivers and passengers regulate their talk around the commitments of the driver to safe and competent driving. In her analysis of car travel Mondada refers to cases of 'convergent' or 'divergent' interaction.



[Participant A is reading the "Time Out" guide. After about one minute she lifts her head, but maintains her gaze on the phone]



A: Oh [*] yeah-------------------
  the gonda [.4]
  gondalean bar i think

[A Turns head down to look at phone again]

B: hhh
A: [*] hhhhh---------------------

[50 seconds reading]

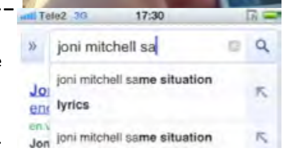
[A looks across to space above B's head]

A: Thats interestinghhh[*]-------
[Raises head and points to poster hung above B's seat]
B: dont ponder it too much you:ll have nightmares
A: hhh
B: just a bit weird
A: very random

Figure 3: Talk and device use - divergent

B: But artists like joni mitchell
A: uuuh
B: are sammy
A: i:s she <u>sa:mi</u>
B: according to wikipedia
A: no ill use the phone [brings out device]
B: who's the other one
B: the other actress the one who used to go out with jack white

[A types Joni m [*], then selects Joni mitchell]-------------------



B: the blonde girl
[A starts new search editing the previous search term, adds 'sa' to the end]
A: how do you write sami [*]----
  is it two a:s
  [types extra a]

B: S a m i [*]------------------

[saami suggested by google]

A: sa:mi i think
[types mi]
B: its just how you say it sami
A: Funny record for your research
A: okkhhhhhha [6.0]
[reads search results]

A: Lapland people lappy [*]------
B: ahha
B: is this are you on wi:wiikipedia sammi people
A: i always thought joni mitchel has big teeth ok well:hhhhhh

Figure 4: Talk and device use - convergent

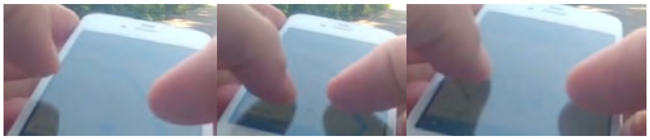
That is, settings where the participants converge in mutual orientation, or alternatively, diverge from that common interaction. Co-participants: "can mutually align [...] and display shared attention towards common objects; they can dis-align as well, in not attending to the other's talk." [23] As she explains, even in cases of 'divergent' attention, parties are still engaged with each other and can time utterances and actions with respect to each other. Figure 3 is one example of 'divergent' talk and device use, in that the participant keeps each involvement relatively separate. The two or so minutes spent reading the webpage are spent mostly without talking to their companion.

In contrast, in figure 4 the search and the conversation are more closely related. Here device use and interaction with others is *convergent*, in that the results of the search at different points are brought into the conversation. Interaction here takes place between the researcher and the participant about the Sami people and a singer. This conversation prompts a web search. As the search is conducted ("Is Joni Mitchell sami?") the conversation does not stop, but rather the search provides topical resources for the conversation. Particularly interesting is the way in which the spelling of sami is approached.

The participant reads out a spelling, at the same time as typing it in. This spelling is confirmed by the companion, but at this point Google makes a suggested autocompletion of the incorrect spelling (from *saa* to *saami*). The tourist then reads out the suggested spelling emphasising an elongated syllable. On conducting the search a conversation ensues about the search before a dismissal of one of the result ("big teeth"). Clearly the search and typing are arranged in concert with the conversation that is taking place. The search responds to and is a source for the conversation. While in this case the device screen does not seem to be directly available to the second conversationalist the participant reads out aspects of the search ("sami", "lapland people", "big teeth"). In this way although her focus of attention is on the device, her involvement in the conversation continues. The use of the device is part of the same activity as the conversation.

At a smaller scale of interaction - with specific user interface events, we can see in a similar way how the device use can converge with interaction. Figure 5 is an example of where a *user-interface* gesture is also an *interactional* gesture. Here the user is talking about an area of the map and as she does this, the participant puts two fingers at each

A: i just check [2.0] which road do we take
B: yes
A: eh yes we will see i think we actually where is it where are we
A: hh [*] e:r [*] e [*]



[Pinches to zoom in on the blue dot]

B: ah ha

[A moves map over to the right, revealing the forest and hiding the suggested route]

A: because I want to go through the forest[*]----------

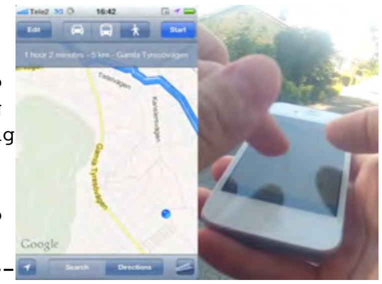


Figure 5: double duty gestures

side of the area referred to, and pinches outwards to zoom in. The pinch enlarges the area of the map where the referred to point is, as well as highlighting the point between the fingers. While the finger movement is away from the desired destination, it does prevent the fingers from obscuring the gestured target.

We would describe this as a form of 'double duty' [30]. In this case the highlighting of an item worked in combination with a user interface gesture. What it also underlines is that interface actions frequently come to have a role to play in the conversation too. At times in our videos, interaction with the mobile device was done as a solitary task - either by a visitor on their own, or on a device turned away from those co-present. Yet devices were also very frequently shared and seen by others. For example, at the end of figure 5 the participant scrolls the map to the right, revealing the forest as she does so. Her reference to the forest is then easily understandable as 'this forest' (she also later points).

As has been extensively documented in studies of interaction around fixed displays, 'awareness' of what is going on amongst co-present interactionists extends to co-present objects and devices in this case, the screen and map which is being manipulated [14]. Using the device, when it is visible to someone else, produces actions that are 'oriented-to phenomena' [12]. That is, those who see the screen can be expected by others to have seen the action and to understand what that action means. Indeed, they often comment on, correct or repair the on screen actions of the device user. Broadly speaking, much of our mobile device use shares this 'double duty' quality - with usage oriented to the device itself (to get it to do what the user desires) but also as something that is seen by others. This is an interactional role of gesture that goes alongside the user interface role.

Occasioned search and search chains

The three data extracts above display interactional features of mobile device use. They are also examples of information search. In figure 3 the participant reads information about areas of Stockholm that they might go and visit. In figure 4, a web search is initiated around Joni Mitchell. Lastly, in figure 5, the participant checks where they are and what route they need to take to go through a forest that is nearby. As we mentioned earlier, much of the mobile device use we captured can broadly be characterised as 'information search' of some sort. Since our participants were making visits to the city this provided a context to much of the searching for information. For example, participants searched for information about what they could do next, the area they were exploring, public transport information about how to get to particular destinations, or information about local cafes, restaurants or shops.

As a time and attention-consuming activity, search is something which requires - at least some of the time - focal attention. Some search we would refer to as pre-meditated search, that is search which takes place before an activity to be part of planning or organising that activity. In earlier work (pre-smartphone) on tourists this was characterised as "pre-visiting" [4], getting information about a place before the actual visit from printed and online sources.

What was striking about the information searching in our study was how it was conducted on the way to or even as part of the activity itself; information would be gathered on the way to a particular part of the city, or even on arrival at the attraction or area of interest. Figure 3 is an example of this, in that the participant is on public transport on the way to the city, and they start to read through a list of 'city highlights' from the Time Out website. The use of resources in the environment is one reason why search may be delayed until the last minute. As you move through an environment you gain information about that place, about the setting, your own mood and what a place is like. Any *in-situ* search can then draw directly upon that information, whereas a search conducted earlier could be premature or waste time. A change in the weather, mood and the like, might affect the planned activity and any subsequent searching that might be undertaken. We are reminded of March's description of decision making in organisations, where decisions are postponed until as late as possible to allow for the collection of maximum amounts of information [21].

This last-minute searching that we observed we titled 'occasioned searches', where searching is triggered by the environment or local events. For example, one visitor (travelling on their own) walks past a postal museum. He searches for information on his iPhone about the museum, reads the short summary and then walks into the museum. After picking up an information sheet (which he quickly discards) and a quick browse of the bookshop he returns to walking along the street. In this case the physical surroundings prompt an online search, which informed the visit. The search was occasioned by the circumstances. In a related way, in figure 4, this search was occasioned by the conversation taking place, where the search term came from a query over a singer's heritage. By "occasioned" here we mean that the details of the particular situation produce the reasons for the activity, and that search is tied into what will be done next. So, for example, coming across an interesting looking tourist attraction might occasion a search. Alternatively, one might come out of a subway and decide to find some lunch - with the search taking place there and then to find a restaurant or cafe. Perhaps surprisingly, many of these searches resulted in failure of some sort, such as not finding the desired establishment, or failing to find the information required, and so participants would fall back on simply waiting to see if they chanced upon what they were looking for (such as a good restaurant), or giving up on the particular activity (such as finding a particular attraction).

Another interesting aspect of information searches was the ways in which the web and the iPhone's inbuilt map application would be switched between in an 'application chain' [3]. This would take the form of searches for information in the web browser and then switching to the inbuilt map to find a place (with the user often typing in the name of a place). This movement between the map and the web took place in part because searches were on names of particular sites and attractions - with the location of those sites being important, but also what those sites were too.

So, for example, in figure 6 a participant sees a church of interest on their map, and then searches for that church on Google. After reading about the church, they then search for the particular area where that church is. This leads them to a shop which is listed at the top of Google's listing, next they click on the map and view the shop on the map. Then they return to Google and visit the shop's website. Finally, they return to the map and study the distance between their current location and the shop (and the searched-for area). The movement of the map between their current location and the shop emulates the route they will later take, and without specifically planning a route they look at the general set of paths they then take to walk down to the searched for area.

MAPS

One of the most used apps in our study was the map application. The success of smartphone map applications, alongside the use of GPS based turn-by-turn driving systems, has affected the sales of paper maps in recent years [28]. While paper maps have advantages of tangibility and size, this is countered by the much more dynamic abilities of mobile device map applications. In our trial, as we mentioned above, searches would often start in the browser (with Google) then move over to the Map application as the location of the searched for item became relevant.

Using electronic maps to navigate is dependent to a large extent on overcoming 'normal natural troubles' [5] - such as inaccurate positions, incorrect maps, unclear maps, miscalculated routes, and problems in specifying destinations, locations and so on. However, these problems are not so much issues that demand redesign, since many of them are practically unavoidable. Maps are always

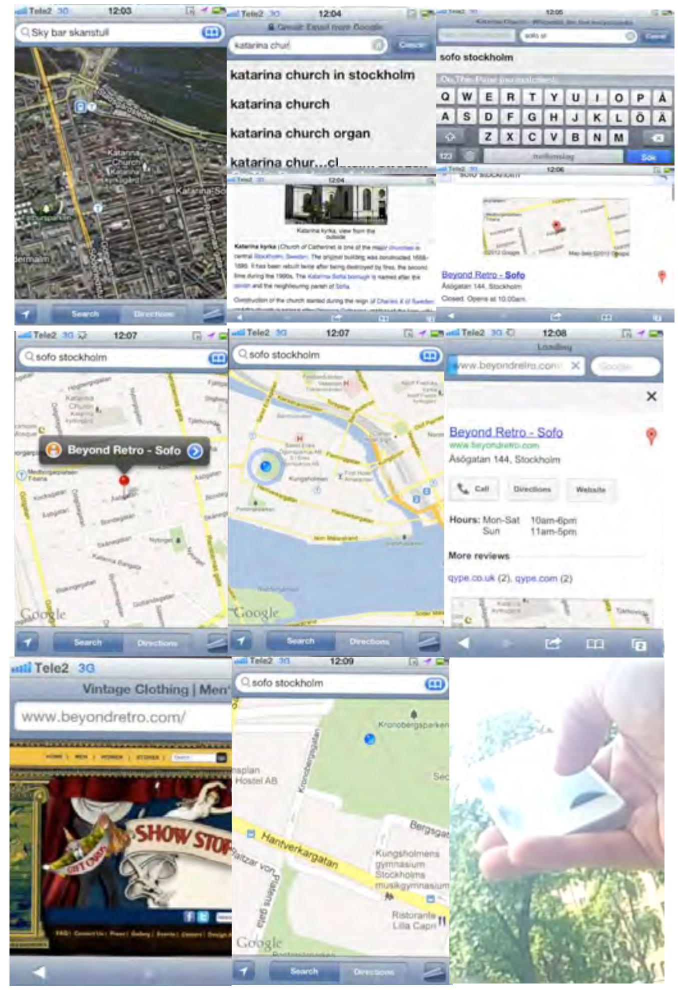


Figure 6: Eight part search chain. User switches between map and browser while sat on a park bench

incomplete and require the wayfinding skills of those present to be successfully followed. For iPhone users public transport information was one potential problem, as the maps app only displayed one subway exit per station, and described trains by route numbers rather than their subway line colours - in one case this even led a participant to try to get a bus rather than enter the subway.

A broader point concerns how the map comes to be used, and in particular, the orientation work that is necessary to move between the map's representation of the local environment and the visible features of the environment. In map studies the classic notion of map orientation concerns the orientation of a paper map to the environment and the map reader [10]. A well established finding is that map users find it easier, and make less mistakes, when they are orientated behind the map, facing the direction of 'up' on the map. This allows for the direct visual alignment of items of the map and the environment - a road to the left then also is on the left on the map [19]. We noticed similar orientation work by our users, with the phone turned to bring the map into orientation with the environment. In figure 7, for example, the phone is held so that the street straight ahead is on the map aligned up.

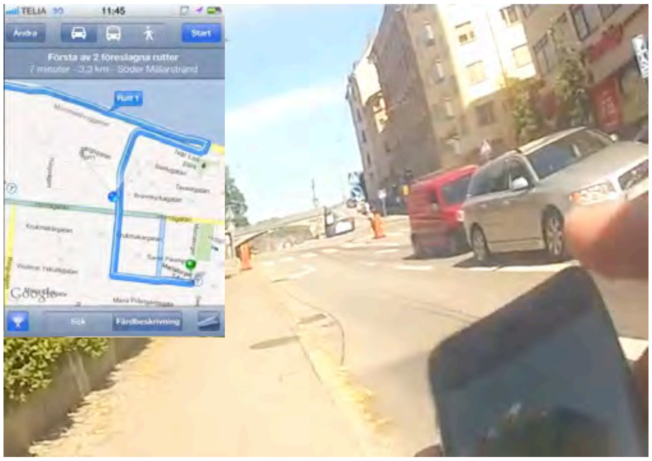
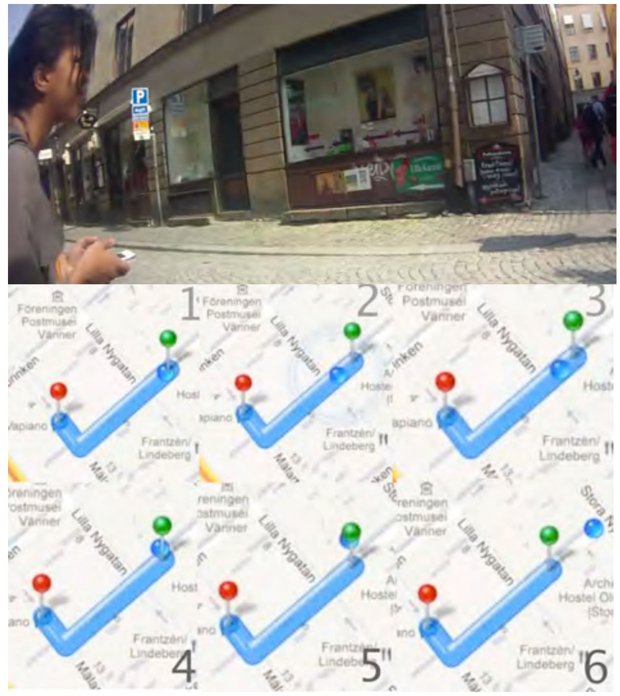


Figure 7: Street aligned with the phone, the blue line and the blue dot

The iPhone that we used in the study feature a 'rotating map' view that should turn the map to align with the environment. Yet the compass functionality is error-prone, and much of the time it would misalign the map with the environment (and users accepted and expected the errors). In some situations (like figure 5) a conversation is taking place around the map and turning the phone to manually align it had advantages, in that particular streets could be accurately (although manually) aligned, and that this could be done relatively quickly so long as one knew the current orientation. Physical rotation also made alignment of the map and environment available to others. With the map oriented to the environment participants could then following a recommended route, with the blue line of the route would stretch up and continue visually with the road being walked along, giving a satisfying visual confirmation of being on the correct route (as in figure 7).

Finding the orientation could however cause problems. Some users chose to discover their orientation themselves. This in part could have been due to problems with the compass, or they were simply unaware that pushing the 'locate' button twice enabled the map rotation. Whatever the rationale, we noticed that some users instead 'walked the blue dot' to try to find their orientation. This entailed walking in one direction then checking to see the new position relevant to the old position. The street walked along could be aligned to the street revealed by the change in position. If the aim was to follow the route a user could also then see if their location was 'following' the blue line.

Unfortunately, one feature of the iPhone's navigation system (at least in iOS 5) disturbed this. To improve the accuracy of the positioning system the iPhone sometimes 'snaps' a users location onto the current route. In many ways this is sensible. Much of the time one will be following this route, and a users location, is more likely to be on that route than elsewhere. Yet when on foot and distances are small and GPS readings inaccurate, this assumption fails. Here, one might accidentally walk tangentially away from the route. In an attempt to 'correct' GPS signals the iPhone would then display the pedestrian walking on the route, giving incorrect confirming of the pedestrians choice of direction. Indeed, when not-driving there are a number of issues that arise in maps use that are less likely to arise



A searches for "cheap eats old town" and finds a comment about a restaurant called vapiano. They type vapiano as a destination into the map and start to walk across the street [pictured, map frame 1]. As they do so, the blue dot (indicating current location) moves along the blue line towards the destination the red pin [2]. After 18 seconds, it changes direction [3] then starts moving away from the destination [4]. After arriving at a street junction the visitor realises they have been walking in the opposite direction from the restaurant. They give up their search for Vapiano and find the closest cafe.

Figure 8: The visitor walks up small side street seemingly following the suggested route.

when driving. Our pedestrians seemed much more likely to deviate from the planned route, than in our earlier studies of driving. Moreover, since pedestrians can walk on a wider variety of terrain than cars (which hopefully remain mostly on roads) plotting a route across a city can be difficult. Shortcuts which are obvious to a pedestrian (such as an open gate, or an unmarked footpath) are not available to the device. Pedestrians might also pick a more desirable route (such as across a park, rather than under a highway) encouraging them to deviate them from the path. City environments are also three dimensional, something difficult to plot or model on a two dimensional map. Suggested pedestrian routes then might often be suboptimal, or downright wrong. If public transport was taken then this also could lead to inaccuracies in the route suggested, since public transport data - although valuable and useful for many of our participants - was also approximate in many ways.

The route snapping technique is particularly problematic for pedestrians when they deviate from the route, in that it provides wrong information, often at a point where accuracy is crucial. Indeed, it can seem to encourage deviations from the route, since these will not initially be rendered. Figure 8 shows one example of a pedestrian walking in the wrong direction, which is seemingly confirmed by the map. Yet after 18 seconds it is revealed that the map has incorrectly plotted their moving position.

When a route was deviated from, at least it was often not a large challenge to navigate in the general direction of the suggested route, or to navigate back to the route. The blue line on the map was still useful even if it was not literally followed (see, for example, figure 5). It is clear that much of the design of the Maps is optimised for driving. This perhaps is wise since the majority of journeys in the US are taken by car. Yet for pedestrians this means that the map tool, while useful, does have shortcomings.

DISCUSSION

We found that the recording methods described here provide a unique and powerful view on interaction around mobile devices. The data we have included highlights a range of different aspects of device use. The news here includes the 'double duty' of interactional gestures, the ways in which information search is occasioned by local activity, and features of using the map such as 'walking the blue dot' This would have been impossible to find without a method that supported this sort of detailed analysis.

There are of course some limitations. As with any complex device the variety of applications and use are determined largely by the user and their situation. Our data then is obviously somewhat constrained in that it focuses on usage on a particular type of day (non-work), in one city. Our reliance on providing a device for the majority of our data collection also constrains considerably the type of application and device usage we collected (alongside our reliance on one model of mobile device). We have no examples of game playing, for example, even though this is one of the most popular categories of applications. With the reliance on mobile cameras, the perspectives and view that is recorded can at times inadequate for analysis. Yet this is balanced by the inclusion of a clear uninterrupted recording of the onscreen activity on the mobile phone. This makes available the 'workplace objects' for analysis in a way not previously possible; something that would be harder to access with traditional ethnographic methods. In turn the use of portable and wearable cameras supports collecting data on interactions that would have been impossible to record using fixed cameras.We have also developed a version of the software that records audio alongside screen interactions, and we hope to capture the device camera as well. We intend to make this software available to other researchers, to assist in future studies.

Mobility *in vivo*

A second concept that our data casts light on is mobility itself. While mobility is a foundation for much technology design, it has had surprisingly little analytic attention within HCI. Our research builds on the work of Luff and Heath - in particular the move away from mobility as mere portability, to 'mobility in interaction' and the myriad ways in which a particular device may or may not fit with situations of ordinary use because of how that device fits with interaction itself. Luff and Heath describe 'micro-mobility' - the ways in which a device would fit with physical bodies and interactional commitments.

Developing this notion further we can see that the form of mobility here goes beyond simple portability. This mobility-in-interaction encompasses the broader range of commitments and activities that we undertake in our daily lives. That mobility is the way in which a drink can be ordered and that transaction interleaved with looking something up on a map. Mobility is what allows information to be looked up as one walks past a museum, immediately seeing electronic sources as one walks into the actual physical museum.

It may be the case that the value of mobile devices come less from their explicit portability, and instead from how we can weave themselves into our ordinary actions and interactions. That is, mobility not simply in terms of physical size but mobility *in vivo.* Mobile devices fit with our lives and actions in ways that conventional technology never could. We would argue that mobility *in vivo* is a topic that still rewards renewed research interest.

CONCLUSIONS

The data here reveals a world of device use that using previous methods is hard to access. It reveals how interaction with smartphones such as the iPhone *in vivo* is threaded through other activities. We documented the differences between convergent and divergent mobility, how gestures on the device are seen by others, the ways in which searching for information is arranged, and lastly we talked at length about how the map is used. Our focus throughout this analysis has been on understanding the relationships between device actions, interactions, physical actions, movement and conversation.

In closing we would argue that the method we have provided here gives us considerable potential to move to a better understanding of device use, and of course how these devices might be better designed. We have resisted detailed design suggestions, instead attempting to build a more substantive understanding of what it is to use a mobile device. Indeed, it may be that the touch and surface computing literature is too focused on developing the latest interaction technique to engage properly with situations of mobile use. Device use is dependent upon and threaded into what goes on around us. It may be that rather than pushing us away from the world around us, our mobile devices are instead just another thread in the complex tapestry of everyday interaction.

ACKNOWLEDGEMENTS

We thank Marek Bell for writing the recording software, the helpful comments on the paper by the anonymous reviewers and our colleagues at mobile-life. The research was made possible by a grant from the Swedish Governmental Agency for Innovation Systems to the Mobile Life VinnExcellence Center, in partnership with Ericsson, Microsoft, Nokia, IKEA and the City of Stockholm.

REFERENCES

1. Barkhuus, L. and Polichar, V. E. Empowerment through seamfulness: smart phones in everyday life. *Personal and Ubiquitous Computing*, 15, 6 (2011), 629-639.
2. Berg, S., Taylor, A., et al. Mobile devices for the next generation: device designs for teenagers. In *Proceedings of CHI 2003, Fort Lauderdale, Florida,* ACM Press (2003).
3. Böhmer, M., Hecht, B., et al. Falling asleep with angry birds, facebook and kindle: a large scale study on mobile application usage. In Proceedings of MobileHCI 2011 (2011),47-56 ACM.
4. Brown, B. and Chalmers, M. Tourism and mobile technology. In *Proceedings of ECSCW 2003,* Klewer Academic Press (2003), 335-355.
5. Brown, B. and Laurier, E. The normal natural troubles of driving with GPS. In Proceedings of CHI 2012 (2012),1621-1630 ACM.
6. Brown, B. and O'hara, K. Place as a practical concern of mobile workers. *Environment and Planning A*, 35, 9 (2003), 1565-1587.
7. Carter, S., Mankoff, J., et al. Momento: support for situated ubicomp experimentation. In *Proceedings of the SIGCHI conference on Human factors in computing systems* (2007),125-134 ACM.
8. Church, K. and Oliver, N. Understanding mobile web and mobile search use in today's dynamic mobile landscape. In Mobile HCI 2011 (2011),67-76 ACM.
9. Froehlich, J., Chen, M. Y., et al. MyExperience: a system for in situ tracing and capturing of user feedback on mobile phones. In MobiSys '07 (2007),57-70 ACM.
10. Golledge, R. G. Human Wayfinding and cognitive maps. In *Wayfinding behavior, cognitive mapping and other spatial processes,* The John Hopkins University Press (1999).
11. Goodwin, C. Notes on story structure and the organization of participation. In *Structures of social action* (1984), 225-246.
12. Greiffenhagen, C. and Watson, R. Visual repairables: analysing the work of repair in human–computer interaction. *Visual Communication*, 8, 1 (2009), 65-90.
13. Heath, C. and Luff, P. *Technology in action*. Cambridge university press, Cambridge, 2000.
14. Heath, C., Svensson, M. S., et al. Configuring awareness. *Computer Supported Cooperative Work (CSCW)*, 11, 3 (2002), 317-347.
15. Heath, C. C. and Luff, P. Collaboration and control: Crisis Management and Multimedia Technology. *CSCW*, 1 (1-2)(1992), 69-94.
16. Hindmarsh, J., Fraser, M., et al. Object-focused interaction in collaborative virtual environments. *ACM Transactions on Computer-Human Interaction (TOCHI)*, 7, 4 (2000), 477-509.
17. Ipsos-MediaCT/Google. *Mobile Internet & Smartphone Adoption: New Insights into Consumer Usage of Mobile Devices, the Shift to Smartphones & the Emergence of Tablets, http://googlemobileads.blogspot.se/2012/01/new-research-global-surge-in-smartphone.html*. 2012.
18. Kamvar, M. and Baluja, S. A large scale study of wireless search behavior: Google mobile search. In *Proceedings of the SIGCHI conference on Human Factors in computing systems* (2006),701-709 ACM.
19. Laurier, E. and Brown, B. Rotating maps and readers: praxiological aspects of alignment and orientation. *Transactions of the Institute of British Geographers*, 33, 2 (2008), 201-216.
20. Luff, P. and Heath, C. Mobility in Collaboration. In *Proceedings of CSCW 1998,* ACM Press (1998), 305-314.
21. March, J. and Sevon, G. Gossip, information and decision making. In *Decisions and Organzations,* Basil Blackwell (1988).
22. Merchant, G. Mobile practices in everyday life: Popular digital technologies and schooling revisited. *British Journal of Educational Technology*, 43, 5 (2012), 770-782.
23. Mondada, L. Talking and driving: multi-activity in the car. *Semiotica*, 2012, 191 (2012), 223-256.
24. Murtagh, G. Seeing the 'rules': preliminary observations on action, interaction and mobile phone use. In *Wireless World,* Springer (2001).
25. Palen, L. and Salzman, M. *Voice-mail diary studies for naturalistic data capture under mobile conditions*. ACM, City, 2002.
26. Palen, L., Salzman, M., et al. Going wireless: behavior & practice of new mobile phone users. In *Proceedings of CSCW '00,* ACM (2000), 201-210.
27. Psathas, G. *Conversation analysis: The study of talk-in-interaction*. Sage Publications, Incorporated, 1994.
28. Rodriguez, B. Amid GPS boom, paper maps hang on, 3rd July 2012. *USA Today*(2012).
29. Scollon, C. N., Kim-Prieto, C., et al. Experience sampling: Promises and pitfalls, strengths and weaknesses. *Journal of Happiness Studies*, 4, 1 (2003), 5-34.
30. Tse, E., Shen, C., et al. How pairs interact over a multimodal digital table. In *Proceedings of the SIGCHI conference on Human factors in computing systems* (2007),215-218 ACM.
31. vom Lehn, D., Heath, C., et al. Exhibiting Interaction: Conduct and Collaboration in Museums and Galleries. *Symbolic Interaction*, 24, 2 (2001), 189-216.
32. Weilenmann, A. and Larsson, C. Local Use and Sharing of Mobile Phones. In *Wireless World: social, cultural and interactional aspects of wireless technology,* Springer Verlag (2001).