

# Eyespy: Supporting Navigation through Play

Marek Bell<sup>1</sup>, Stuart Reeves<sup>1</sup>, Barry Brown<sup>2</sup>, Scott Sherwood<sup>1</sup>,  
Donny MacMillan<sup>1</sup>, John Ferguson<sup>1</sup> and Matthew Chalmers<sup>1</sup>

<sup>1</sup>Department of Computing Science

University of Glasgow, UK

{marek, stuartr, sherwood, donny, john,  
matthew}@dcs.gla.ac.uk

<sup>2</sup>Department of Communication

University of California, San Diego, USA

barry@ucsd.edu

## ABSTRACT

This paper demonstrates how useful content can be generated as a by-product of an enjoyable mobile multiplayer game. In *EyeSpy*, players tag geographic locations with photos or text. By locating the places in which other players' tags were created and 'confirming' them, players earn points for themselves and verify the tags' locations. As a side effect of game-play, *EyeSpy* produces a collection of *recognisable* and *findable* geographic details, in the form of photographs and text tags, that can be repurposed to support navigation tasks. Two user trials of the game successfully produced an archive of geo-located photographs and tags, and in a follow-up experiment we compared performance in a navigation task using photographs from the game, with geo-referenced photos collected from the *Flickr* website. Our experiences with *EyeSpy* support reflection upon the design challenges presented by 'human computation' and the production of usable by-products through mobile game-play.

## Author keywords

Human computation, mobile multiplayer games, mobile photography, navigation, RF fingerprinting.

## ACM Classification Keywords

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

## INTRODUCTION

In recent years, location technologies, such as GPS, have enabled a range of new mobile applications. For example, digital cameras can now automatically tag photographs with their location (e.g., Eye-Fi ([www.eyefi.com](http://www.eyefi.com))), allowing images to be browsed and arranged geographically. A number of photography websites, such as Flickr ([www.flickr.com](http://www.flickr.com)) and Panoramio ([www.panoramio.com](http://www.panoramio.com)), to provide large collections of publicly available, accurately geo-referenced images.

One potential use of these images is for supporting navigation, in that maps can be labelled with corresponding

*in situ* perspectives of city locations [2], addressing some of the orientation problems commonly found in the use of maps. Indeed, more broadly, collections of geo-tagged information, such as photos, have the potential to transform the flat 'official' view usually represented on maps. Combining photographs and location technology also supports new interactions that extend everyday photographic practices, for example, games that combine photography and location, as demonstrated by schoolchildren's inventive appropriation of location and camera phone technology for enjoyment [17], and geo-caching [16], as well as games that involve other forms of mobile photography (e.g., the photographic 'mission' oriented play of [9]).

In this paper we explore both of these opportunities, describing our experiences with *EyeSpy*, an application designed to generate navigation-ready photos and labels from a mobile multiplayer game. Building on 'human computation' research, such as the ESP Game [20], *EyeSpy* is mobile game with 'useful by-products' in which the quality of the by-products are verified in-game. The game involves players taking photographs and entering text labels for local landmarks. The game dynamic encourages 'good' photos of 'what everybody can find' in a particular area, giving players extra points for images that can be found by other players. A side effect of play is the production of a corpus of photos and text labels that can easily located.

*EyeSpy* thus generated collections of geo-located images, but unlike websites such as Flickr, these are photographs specifically suitable for navigation (or at least are easily findable). As a game, *EyeSpy* also explores how photography can be seen not only as a hobby or personal interest, but manipulated into new forms of leisure—such as photography games, or as a method of bringing a social group together. The version of *EyeSpy* trialled for this paper can be played without any specialised hardware and in most locations, making use of wifi triangulation to geographically locate players and their text or photo tags.

In two user trials of the game, collectively involving 18 players, we built up a corpus of 257 images and 196 text tags in one geographically bounded area. We then tested the photographic output of these games in a simple navigation task, comparing the photographs to geo-located images stored on Flickr. This subsequent test demonstrates that the

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

CHI 2009, April 4–9, 2009, Boston, Massachusetts, USA.  
Copyright 2009 ACM 978-1-60558-246-7/09/04...\$5.00.

game produced photographs that were markedly better than Flickr images for navigation albeit, as we note, with some limitations.

We start by discussing prior work that influenced the development of EyeSpy, followed by an overview of the EyeSpy game. We then explore how players approached the game, and the different strategies they used to take photographs and tags. Lastly, we discuss how the by-products of the game fared in two simple experiments.

### **PREVIOUS WORK**

The primary aim of EyeSpy was to extend so-called ‘human computation’ [20, 8] to a mobile setting. This approach makes use of humans’ abilities to do what computers cannot. Human computation’s best-known examples are web-based games that have by-products in the form of beneficial data, such as the ESP Game [20]—a game now launched commercially as Google Image Labeler ([images.google.com/imagelabeler](http://images.google.com/imagelabeler)). The ESP Game involved players attempting to match descriptive tags for images, resulting in the rapid collection of annotations for large numbers of images. Verification of the results is achieved in part because players are unknown to each other and the game mechanic makes it unlikely that players can collude or cheat to insert inappropriate tags.

This approach has been partially applied in mobile settings. For example, the Treasure game [1] was designed to create maps of wifi coverage in an urban area as a by-product of game-play. Treasure’s by-product takes advantage of human movement but does not involve human computation as such, because wifi strength is an objective measurement straightforward to achieve with computers. The CityExplorer game [13] involves exploiting commonsense, local knowledge. In mobile play, urban areas were tagged with categories such as ‘church’ and ‘beer garden’. Verification of the tags was done in a later web-based phase, in which players judged the correctness of one another’s tags via a web-based interface—a process that Matyas et al. reported that players found “cumbersome” [14]. EyeSpy builds upon such prior work in that we exploit participants’ commonsense local knowledge to produce a useful image set, and incorporate a verification mechanism as a fundamental and ongoing part of the mobile game.

EyeSpy also draws on recent innovations in photography and photo taking, a long-standing interest in HCI and related fields. The growth of camera phones, and new positioning technology (e.g., GPS) to geo-locate images and share them with others has resulted in a range of new applications (e.g., Yahoo’s Zurpher and IDEixis [15, 19]).

Drawing on these developments, we sought to explore the potential for a pervasive game based around geo-located photography. Pervasive games take place over a geographical area, and usually over a long period of time, attempting to break the user experience away from the desktop or handheld and push it further into the everyday world. One early pervasive game, Can You See Me Now?

(CYSMN) [4], was played on city streets, combining online and physically present players. Over time this genre has grown in a number of interesting directions. In the Feeding Yoshi game [3], the goal was not only to expand the area on which the game was played, but also to weave the game into players’ everyday lives—with game-play designed so that it could be interspersed with everyday life and work over a period of weeks, rather than demanding concentrated use for minutes or hours. Feeding Yoshi enabled play when players had free moments, but also made use of players’ everyday movements as a key game dynamic. Thus, an important aspect to pervasive gaming is the way in which players’ lives, knowledge and location become key parts of the game. In CYSMN, local knowledge of where to run was key to players’ success; in Feeding Yoshi it was knowledge of areas expected to have a high density of wifi access points.

### **A GAME WITH USEFUL BY-PRODUCTS: EYESPY**

This involvement of local knowledge as a key part of game-play is explored further in EyeSpy. In designing EyeSpy, we produced a simple game that could both be played dynamically over a long period of time over a city’s streets (like Feeding Yoshi and CYSMN), and made use of geo-referenced photographs as a key part of the game, producing beneficial by-products (like the ESP Game).

In EyeSpy, players take photographs that are shared with other players, who then have to find where those photographs were taken in order to confirm or validate them. This validation helps address the new possibilities of cheating in a mobile environment when compared to a web-based system like the ESP Game. Points are scored by players for both confirming others’ photographs, but also for producing photographs that were popular for other players to confirm. Players submitting a photograph are therefore concerned with authoring a picture that is likely to be confirmed by as many other players as possible, producing images that are easy to recognise and find by others. As an alternative to taking photographs, players can write short ‘text tags’, which again are confirmed by other players by physically going to the places where the tags were created.

### **Game design**

The key design goal of EyeSpy was to produce a game that would generate geo-referenced photographs and tags suitable for map annotation. Orientating maps to the environment is a challenge for many map users, and literally just finding where you are on a map as you stand on a street can be a challenging task [12]. Providing a photograph on a map at the right place could greatly assist this task, as well as overlaying the map with the ‘texture’ of the area [6].

While there are a number of collections of geo-referenced photographs already available (such as Flickr, Google StreetView and the like) one key problem with these sources is knowing which photographs are usable for navigation. For instance, Flickr photographs are taken and uploaded for a wide variety of reasons (e.g., art, amusement, emotion) and while these are an important part of the value of sharing

photographs, it can conflict with using photographs for navigational purposes. Even amongst repositories that are more utilitarian in design, such as the photographs collected by Google StreetView, one is left with the problem of selecting which photographs to use from a stream of millions. While selecting appropriate photographs by hand is practical for small areas, overall it depends upon employing local knowledge of what are good and familiar local landmarks, and this is difficult to do over a large area.

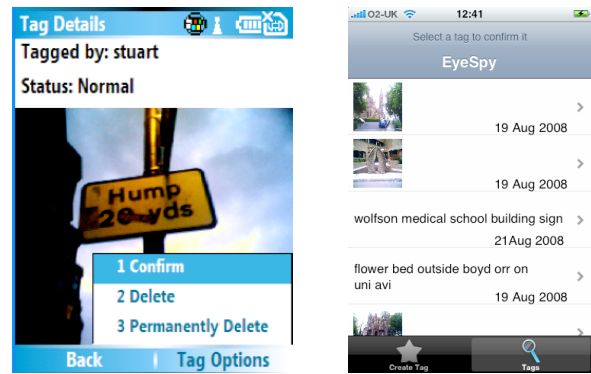
Therefore, our key design goal in EyeSpy was to reward players for producing geo-referenced photographs of good landmarks or easily found objects. In the eventual game design this became geo-located ‘tags’ that could be easily found by other players: players score points by having their tags confirmed. Players had no control over who could confirm these tags. While the game can be played in small groups of friends, we also designed the game so that it could be played amongst groups of strangers. Apart from blocking collusion among players, we hoped that this lack of control would mean that players would take photographs and write labels that would be sufficiently generic to be findable by any other player. We also reasoned that these photographs would potentially be suitable for navigation since, if they could be easily found by other locals, they would potentially also be easily findable by those unfamiliar to the area.

In the game, at the start of each day, every player receives a set of 10 tags (five photo and five text tags) to confirm (see Figure 1, right). This set is randomly chosen and anonymously presented. Players then get points for both confirming these tags (Figure 1, left), and also for creating new photo or text tags. When a player’s tag is confirmed by another, the author is notified and his or her score is increased.

### Technology

Technically, EyeSpy runs on commodity hardware mobile phones, with wifi being the only ‘high level’ feature needed. Our trial system ran on iMate-SP5 phones using the Windows Mobile platform (the most recent version of EyeSpy is for the Apple iPhone).

One of the goals of EyeSpy was to allow play at any urban location, including indoors, so that by-products could be generated for any desired area. However, in selecting a method of locating users we were constrained by the power, storage and processing capabilities of the devices used. We elected to use radio frequency (RF) fingerprints to match the locations of tags users created in the game. Not only did this prove to be an extremely efficient technique on the client devices, it allows for subsequent rapid matching of all the uploaded tags. Through detection of the unique IDs of local RF beacons (in EyeSpy, 802.11 wifi access points, but could be GSM cell antennae) and signal strengths these beacon transmissions can be used to generate a unique pattern, or



**Figure 1: The Windows Mobile (left) and iPhone (right) versions of EyeSpy. A photo is about to be confirmed (left), and (right) a list of photos and text tags to confirm is shown.**

fingerprint, which characterises a particular location. Once this fingerprint is stored a subsequent scan may be used to determine whether the current fingerprint overlaps with the recorded fingerprint, thus ascertaining if the device is at the same location. Using wifi access points gave the game a quite high granularity and accuracy of locating tags and photos, as 802.11 beacons typically have a maximum range of 100m—much lower than GSM—and in most cities are generally distributed more densely. In our trial, we found that the average EyeSpy fingerprint was constructed from data of 7.99 access points. We required at least a 50% overlap before the current scan was said to match a fingerprint, resulting in scans being matched within approximately a 5-20m range.

In EyeSpy we store fingerprint data—access point MAC addresses and signal strengths—along with text and photo tags. These fingerprints are thus tied to images and textual descriptions on the client device, however we note that the fingerprints are not converted to geographic coordinates. Geo-referencing photographs is becoming increasingly popular, however there are some issues with the use of GPS. Where GPS hardware is available, it often requires time to ‘warm up’ before it can get a position fix—often several minutes—which does not provide a suitable match with the timeframe of taking a photograph—often seconds. Wifi positioning is generally faster, e.g., fingerprinting eight hotspots takes less than 1s whereas GPS from a cold-start can take minutes. In addition, when using GPS hardware, the device must be powered constantly, thus draining battery power, or be powered on after the photograph is taken, which may lead to the first GPS fix several minutes later being inaccurately used as the position of the photograph. Wifi also favours built-up areas where GPS may encounter considerable problems (e.g., ‘shadows’). Our coupling of fingerprints and images rather than explicitly geo-coding these images can provide a more suitable method to position photographs on mobile devices. However, we note that GPS does favour a wider availability, and there will always be different areas of applicability for wifi and GPS positioning.

## TESTING EYESPY

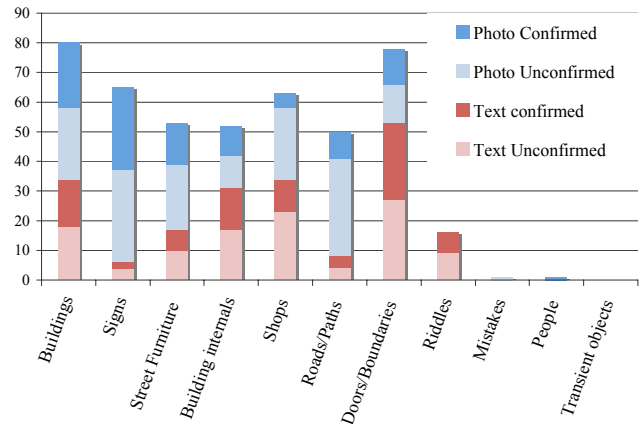
In order to test the game, we ran a trial involving 18 (6 female and 12 male) participants over two separate rounds. In the first round nine participants played for one week, with the subsequent nine participants playing a second round lasting two weeks. In the first round we focused on encouraging as much play as possible, seeding the game with our own photographic and text tags at the start of play. In the second round, the game was both not seeded and played for twice as long. Players in the first trial were paid £10 to play, with the winner getting a further £10. In the second trial, players were paid £10 per week, with the winner being awarded a further £20.

In the first round, players were drawn from Computer Science undergraduates, who knew one another before starting the trial and who had existing social ties. The second round involved a more mixed group of seven non-Computer Science students plus two non-student participants. Eight of the players in the group were acquainted with each other, although this group on the whole did not have strong social bonds. The players in both trials worked, studied or lived in the area of the city around our university. This acted as a natural limit on the ‘game area’. In addition we asked players to restrict their play to around this area, so as to prevent the game becoming too difficult to play. However, players did at times make text and photographic tags outside this area.

After the trial players were interviewed, with the interview transcripts then coded and analysed for key themes. (Note that these themes were derived jointly from this and our Flickr photoset, which is mentioned later on in this paper.) We paid special attention to the reported motivations of players, the different game styles and strategies that players adopted, and the information shared and relationships between players both through the game and outside of the



**Figure 2:** A portion of the photos collected during the two trials of EyeSpy (area shown is approx. 400m<sup>2</sup>). Regions in gray are streets and open areas inside the University of Glasgow.



**Figure 3:** Tags taken by players categorised by photo or text content. Some tags are in more than one category as they contain multiple types of content.

game. Lastly, we looked at where players went and their relationship with those places in playing the game. In particular we were interested in how the game dynamic developed and how players oriented to the rules of the game in producing their photo and text tags.

### Tags created and confirmed by players

Broadly, the players created a mix of photo and text tags; out of 453 tags overall there were 257 photos (57%) and 196 text tags (43%). Players produced on average 25 tags during the game: 14 photos and 11 text tags. Figure 2 shows a portion of the image tags generated by players from both rounds. As might be expected, the tags were mainly of geographically prominent landmarks—statues, street corners, shops and the like. Figure 3 shows a categorisation of the tags according to their context. After buildings, the second most frequent tag was doors/boundaries (i.e., entrances to buildings or boundaries between different places). Interestingly, players also made text tags involving simple riddles, such as creating the tag ‘music to my ears on uni gardens’ (rather than explicitly stating the tag’s location outside the Music Department, on the street called University Gardens). Crucially, there was only one photo tag involving people as the main focus and no tags taken involving ‘transient’ objects in the environment (such as cars).

Tags submitted by players were also confirmed in the game by other players—a player must physically be where a tag was made in order to be able to confirm it. 33% of player-generated tags (average of both rounds) were confirmed, however some tags were confirmed multiple times. In the first round, 21% of player-generated tags were confirmed, although if we include the ‘seed’ tags that we used to start the game this rises to 43% as nearly all the seed tags were confirmed by players. In the second trial, which involved no seeding, 40% of tags were confirmed, suggesting that the confirmation rate increases with more prolonged play.

### Player motivations in EyeSpy

One of our first concerns with EyeSpy was how well the experience worked as a game. As can be inferred by the large number of tags entered by each player, the game did have

some success as an experience, with players reporting that they “enjoyed” the game, that it was “fun”, as well as “easy to use”.

The game presented a number of different motivations for players. Firstly, the players’ scores were visible within the phone application. This proved to be a significant motivation for some players, echoing results from experiences with the ESP Game. Indeed, some players contrasted this motivation with financial rewards for playing the game, such as prizes or payment. Websites that attempt to distribute small tasks over the Internet, such as Mechanical Turk ([www.mturk.com](http://www.mturk.com)), frequently work on the basis of small financial rewards. Yet for our players the ability to compete—and win—against their friends and even complete strangers, was more than adequate motivation, in spite of being a small sum paid to participate (which obviously did provide some level of motivation in and of itself). As a pair of players stated:

You’ve got two very competitive people here, and you’ve got someone who’s first, and we want to move up to first position

A second motivation came from the interaction that grew between players around the photographs. On the whole, forums for social interaction in the game were fairly limited, since it was only tags and scores that were shared between players. Moreover, in the second longer round of the game the majority of players only knew one or two others well. Yet despite this players talked about how even this narrow channel did provide awareness of other players, and in turn more motivation for play. Players talked about being connected to other players in that they created tags of similar landmarks. One player mentioned “walking in the footsteps” of other players, confirming their tags, but also creating tags that were in response to previous tags taken.

Yet the game did lack much in the way of other communication channels to support sociability. A number of players expressed disappointment that they could not easily find the other players, and most expressed a concern about noticing others playing the game:

When I was walking around taking photos I was wondering if I’d run into anyone else with the same phone, or if somebody would spot me with the phone and be like ‘ha’.

This following at a distance was deliberately played upon by the game’s name—EyeSpy. As with many UbiComp experiences that involve tracking, EyeSpy raises a range of issues about privacy, tracking, self-monitoring and the like. While a game played as part of a trial is perhaps an insufficient test of these issues, it is worth noting that none of the players mentioned these concerns while playing the game (perhaps because of the lightweight connections between players). It would be difficult, for example, to be able to gain specific, rather than very general, location information about other players. Photos and text tags were shared without any identifying information, and were selected at random from the pool of contributed tags.

A further issue related to the motivation was how the game encouraged new interactions with the environment. Participants frequently mentioned how the game provided opportunities to explore new areas, as well as experiencing well-known locations in a new light. Players also noted the health benefits of the play of the game increasing the amount they walked around; one stated:

I remember walking round for ages thinking, you know, this is probably good for me in some way [...] I like the exploratory part [...] it reminded me of when, you know you go to another country, and you’re wandering round for ages

### **Player strategies, designing for navigation**

This notion of ‘wandering and exploring’ brought us to consider how players oriented to one another’s movements when taking part in the trial. Like other pervasive games, in order to play EyeSpy players must leave their homes or workplaces, and travel around the streets to both make and confirm others’ tags. However, the act of creating a text tag or taking a photograph involves the player in a specific interaction with their environment. This draws on players’ local knowledge—selecting where to author a tag from all the potential photographs and textual descriptions that might be made.

Moreover, to be successful in the game players needed to consider what other players’ local knowledge is likely to be. Co-players need to be able to find a given players’ tags for that player to receive confirmation points. For photo tags it is important that the photographed object is recognisable and can be quickly found, at a glance, from all the potential places in the game area. For text tags it is in turn important that the text can be used to locate an area quickly and with sufficient accuracy to be registered by the game as the same area. To be successful at the game players therefore have to consider what a ‘general’ player might know about the area the game is being played in.

For players, the key challenge of the game thus came from this problem of how and where to author tags so as to be successful, i.e., to increase their chances of winning the game. That is, they made tags that others could quickly identify and be willing to locate in order to confirm. Authoring such ‘good tags’ therefore required *recipient design* [18]. Players creating tags took into account the perceived behaviour of other players. This ranged from their knowledge of the area, to their expected route and even to their social role. One participant, who knew one other player socially but not the others, tried to author his tags for what he thought the other players (identified as ‘students’) would be able to find and confirm:

I thought that a lot of places I didn’t know where you were so I’ll go for the obvious targets ... obviously a lot of the places I didn’t know where they were because they were uni names so I thought I’d go for the obvious targets

Players consistently referred to the importance of making a photo tag “recognisable”, “identifiable” or “obvious” using what they considered to be a “landmark” or “central” place.



What constituted *recognisability* for players revealed a concern for the navigational experiences of other players. For instance, at one point a player decided to tag a “gargoyle thing that’s got a unicorn” which she considered to be “quirky”, but then changed her mind since “maybe other people won’t know about it”. This concern is also revealed in another player’s comment about making a ‘good tag’: “if [other players] can recognise [a tag], that’s enough”. Producing something recognisable would mean that other players would “know exactly where [tags] are when they see them”.

Of course, some tags would not be instantly recognised by players—recognisability involved a design for *findability*. Players, in considering how findable a tag would be, often reported hypothesising over how other players would go about navigating to a tag. One player reported changing her play strategy in order to achieve this, at first tagging “random places” and then starting to consider questions such as “where would [people] walk?” denoting her attention to the navigation practices that other players she assumed would come to engage in when seeking out her tag.

Finally, what determined a ‘good tag’ also sometimes depended upon a relationship to other tags. In creating tags that were “as easy as possible for people” to confirm, one key technique simply involved putting tags spatially close together and “think[ing] about how people were working” in order to decide where to place them. Players in this way produced a trail of tags that could be created in the course of one walk, but also that a prospective fellow player could walk along confirming multiple tags. Unfortunately, because of the game dynamic—only giving a subset of tags to each player to confirm—it would be unlikely that they would receive more than one tag on any trail negating this potentially beneficial strategy.

Players consistently oriented to the concerns of recognisability and findability when acquiring photos in order to make ‘good tags’ for the purposes of navigation within the game. Crucially this involved players designing their tags according to ‘what anyone knows’—i.e., shared local geographic knowledge [16]—and the presumed activities and orientation of these fellow players to the game (see [7]). As one player commented, for example, “you could tag the Mitchell [Building] because everyone knows where it is, but who’s going to be [bothered] to tramp across town?” This thought is important when considering how the photos might provide useful navigational tools outside of the game since they are exploited within the game as a form of ‘pictorial instruction’.

In turn, confirming a tag demanded some detective work in finding where the tag was, going to that location, and then attempting to position oneself in the same site that the tag was made. Players pointed out that this was easier for photo tags, since with text tags there was typically a much greater ambiguity about exact location. This is perhaps one reason

for there being a greater number of photo tags when compared to text tags generated in play.

In this way, confirmation of tags also reveals a more detailed level of findability: when a player had successfully located the general area of a tag, regardless of whether they had experienced it as recognisable straight-away or instead needed to search it out, the player then had to align their phone’s current fingerprint with the fingerprint of the tag. This was done routinely by players; they located ‘exact spots’ by aligning themselves as demonstrated in the photographs. For instance, a player reported being “sure [they] had the right place” given that it was “exactly the same as it is on the picture”. Players anticipated others’ actions, and chose orientations and alignments to their photos to make confirmation easier for others. As one player stated, it would be “easier for [another player] to figure out where I was standing”. This aspect is key when considering the navigational qualities of the photos taken, however we note that this is less the case for text tags, although players did sometimes design them for findability through explicit instruction, such as “boyd orr building facing qm [union building]”.

#### **Game area saturation**

In the dynamic and flexible way tags could be created, however, the game did have some shortcomings. In particular *saturation* could be achieved in a given area quite quickly when most ‘obvious’ landmarks or easily findable areas had been photographed or tagged. This was due to the players’ orientation to taking tags for a generalised ‘co-players’ route’ as well as orienting towards designing for findability, resulting in players gradually being unable to resolve this orientation for increasingly more ‘obscure’ landmarks. Indeed, the winning player achieved a win through finding and tagging more recognisable landmarks within those everyday routes, rather than by taking pictures of increasingly obscure places. Saturation also occurred thanks to a low effort threshold, brought about by the game being played as the part of everyday life; players were only willing to sacrifice a limited amount of time going outside of their daily routine.

However, some players did attempt to overcome saturation via increased creativity in constructing tags within highly saturated areas. One player mentioned, for example,

trying to take a photo of the same monument [which was already tagged] but from an angle that was a wee bit more abstract; I was almost forced to be artistic in the way that I took it because I wasn’t really near anything obvious, and I knew there was quite a lot of different signals in the area, so I thought if I stand next to this tree in a certain way that people would be able to figure out which tree it was [and therefore be able to confirm the tag]

In addition this further illustrates the concern for recognisable and findable tags that players oriented to, and the lengths they would go to overcome this problem when faced with difficult, saturated areas.

## TESTING GAME BY-PRODUCTS

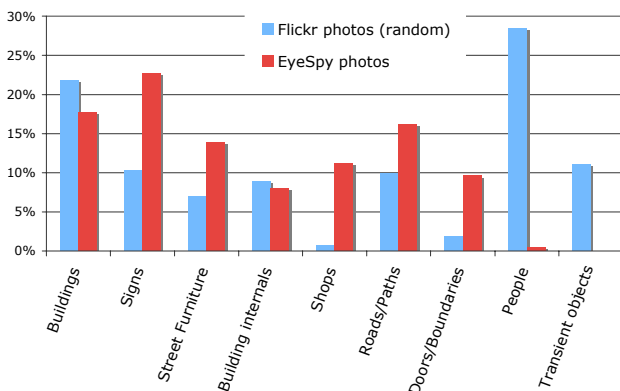
So far we have discussed the qualities of the tags produced by EyeSpy players—particularly images—in the context of their use within the game itself. However, fully assessing the quality of the by-products also requires examining the potential for their use outside the game as stand-alone resources. This final section of the paper explores the results of further experiments that we conducted, focusing on the images generated by EyeSpy, in order to validate our claims about the photos’ recognisability, findability and use for navigation. We note, however, that recognisability and findability are of course only two components of navigation, and in this sense our experiment was configured to test only these two aspects of navigational practices.

We began to consider how our set of generated photos would compare with other collections that could conceivably be used in a location-based service delivering images to users. Of course, there are many different uses and contexts in which a navigation system could be employed (as discussed in [11]). Each different usage could potentially require subtly different types of images. Broadly, however, all images used for navigation share the ability to be quickly identified by individuals *in situ*. The more obscure an image is, and the longer it takes for an individual to visually link a photo with the scene they are navigating, the less useful it is likely to be.

In order to test our by-products, we conducted a series of tests comparisons between our photo set and sets derived from geo-located images available on the Flickr website. Flickr was chosen firstly in response to earlier work suggesting the use of Flickr images for navigation [2], but also because of its worldwide coverage. Whilst our EyeSpy-generated image set was small in comparison, it was relatively dense along particular routes (see Figure 2) and, of the image sites we looked at, only the Flickr website had a comparable photo density, i.e., would permit a fair test.

### Image retrieval and coding

Firstly we needed geographic locations for each image in our EyeSpy set. To do this we resolved the wifi fingerprints to their corresponding GPS locations by aggregating the previously war-driven (i.e., GPS) locations of individual



**Figure 4: Photos from Flickr and EyeSpy sharing similar geographic locations, categorised by content (‘riddles’ and ‘mistakes’ omitted for clarity).**

access points. Based on the location of each of our images, we downloaded a random image chosen from the ten geographically closest images available on Flickr (i.e., 257 in total), which of course themselves contained GPS metadata. We analysed this set (in tandem with the EyeSpy set) in order to derive the key categories mentioned earlier in this paper. Figure 4 illustrates how the Flickr images compared with our own set of EyeSpy photos. In comparing the two image sets’ ‘signatures’ according to our categorisation scheme, we can see that randomly selected proximal Flickr photos contained significantly higher proportions of people and transient objects. Further to this, the EyeSpy image set contained a higher proportion of signs, shops and doors/boundaries, revealing how the game produced greater numbers of photographs of images likely to be more navigationally useful in terms of our criteria of recognisability and findability.

### Navigating with the image sets

Our first experiment involved the construction of two routes within the game area (of length 530m and 800m respectively). For each route, 16 of the closest EyeSpy and Flickr photos to the route were gathered, resulting in four sets of photos:  $R_{1E}$ ,  $R_{1F}$ ,  $R_{2E}$  and  $R_{2F}$ . We employed a two factor experimental design with replication in order to test the various permutations of routes. 10 participants (7 female, 3 male) with limited or no knowledge of the game area were recruited, and each was provided with a simplified map of the area (with street names erased) which had approximately-circled indications of the locations of the photographs for that particular route. Note that, in order to avoid bias, these locations were approximate enough to geographically cover the true locations of corresponding photos spanning the Flickr and EyeSpy versions of the routes (see Figure 5). Thus we only constructed two maps.

Participants were also provided with the photos relevant for their routes. Each participant was then sent out individually to walk a selected pair of Flickr and EyeSpy routes, the objective being to locate where they thought the photos had been taken from. Participants were recommended to spend no more than half an hour to complete each route, and were told that the order in which photos were confirmed did not matter.

As a result of this experiment, we found that participants were able to identify the locations of 95% of EyeSpy photos included in the two routes (91% and 99% for  $R_{1E}$  and  $R_{2E}$  respectively) compared to 49% of the Flickr photos (54% and 45% for  $R_{1F}$  and  $R_{2F}$  respectively). On average participants took 25 minutes 11 seconds (s.d. 8:09) to complete the Flickr routes, whereas on average it took 17 minutes 25 seconds for EyeSpy routes (s.d. 6:16). These results suggested that photos generated from EyeSpy were more frequently and rapidly located than those from our Flickr set.

In order to test these hypotheses, the significance of the effects of each condition were investigated. The statistical



**Figure 5: A route map with approximate image locations indicated (centre); EyeSpy (left) and Flickr (right) images for location 11.**

analysis used is a standard two factor ANOVA analysis, based on the critical values of the F distribution, with  $\alpha=0.05$ . The ANOVA shows there are significant differences in the data between both in terms of success in locating images ( $F=111.82 > F(2, 80) = 3.87$ ) and times taken ( $F=5.42 > F(2, 5) = 4.49$ ). Tukey's pairwise HSD analysis showed that the average number of successes in locating EyeSpy photos, as well as the time taken to do so, was significantly greater than for Flickr photos. There were no interactions between the different routes.

### Selecting 'navigable' images

For our second experiment we recruited a further 16 participants, presenting each with a randomly ordered collection of all the photos from the previous experiment (i.e., 32 Flickr and 32 EyeSpy photos—64 in total). Each was then asked to indicate which photos he/she felt would be "most appropriate and useful for tourists navigating around a city they were unfamiliar with".

We found that 61% of the 32 EyeSpy photos were chosen as appropriate for navigation, whereas only 20% of the 32 Flickr photos were selected, further confirming that our EyeSpy-generated photos were preferable for navigation.

### Text tags as descriptions for images

It was originally intended that the game's design would result in successful photographic by-products that, in combination with text tags, offered the possibility of associating those images with relevant search terms culled from nearby text tags. Although we mostly focused on the images produced by EyeSpy, we also examined the textual descriptions generated by text tagging in order to discover whether they could be effectively reused in this way. However, we found that the way in which text tags were crafted by players, such as the creation of riddle tags and lack of high enough density to provide appropriate search terms, resulted in less than 20% of the closest text tags to images providing relevant descriptions, thus confounding this design aim.

## DISCUSSION

In crafting easy-to-recognise, easy-to-find tags that leverage local geographic knowledge, our initial experimental results suggest that players generated a more focussed set of images than could be found from geo-tagged equivalents drawn from websites like Flickr. Although limited in size, the player-crafted set of photographs generated by EyeSpy thus appears to provide a high quality navigational guide to the area in which the game was played.

However, although EyeSpy's design encourages the production of by-products that have potential reuses for navigation, as well as creating a largely enjoyable experience for players, our evaluation of EyeSpy also reveals some design tensions.

### Game design and by-products

In running the game we were particularly interested in how the game dynamic and the need for 'good' by-products would complement or conflict with one another. That is to say, what might make an enjoyable game experience might not make for good tags and photos. Like many designs, here we were faced with multiple constraints or requirements, in that the overall design had to serve two purposes concurrently—a good user experience within the game and useful by-products coming out of the game. Although not necessarily in conflict, or the result of a simplistic trade-off arrangement, by-products and game rules must be carefully balanced.

One feature that makes design of a game with by-products unusual stems from the way that, like any game or user experience design, successful design involves considering how formal rules will be used and interpreted in practice. A game designed to create by-products harnesses the enjoyment, intelligence and creativity of players. We need players to enjoy such games so that they are motivated to make by-products, but we also want them to play only in ways that create what we consider to be useful by-products. We want them to be creative but not 'too' creative; that is, players' enjoyment and engagement may encourage them to find ways to play that help them win, or help them enjoy the game more, however do not create useful by-products. EyeSpy's rules orient players toward a strong concern for two aspects of navigation, recognisability and findability—which is in accord with our repurposing those images for navigation—but nevertheless the rules encourage rather than enforce such an orientation (see [5]).

Players' practical engagement with the rules, and even the language we used in order to introduce and 'frame' the game (e.g., using the name 'EyeSpy' confused some players initially due to the children's game 'eye spy'), configured certain expectations about the style of play. By and large, this oriented players toward the production of 'good' photographic by-products but, as mentioned earlier, some players also created 'riddles' within their text tags that required co-players to engage in some detective work in order to confirm them. Although these tags might have well been more enjoyable for players both in creating them and discovering how to confirm them, they were less useful as by-products, and lacked the recognisability and findability that were touchstones for well-constructed, 'good tags' for most EyeSpy players. While it may have been more fun for some to engage with this form of tag, the work required of players in locating them conflicted with the need to maximise tag confirmations, which in turn meant that riddles did not earn many points for the creators. Thus, the rules of EyeSpy



to some extent discouraged such creative and playful activities. This serves as an example of game rules encouraging one potential style of play, rather than enforcing it; riddles are technically feasible but lose out to a style based on more straightforward recognisability. One can imagine, though, that a determined enough group of players or a slight shift in game rules might lead to a different game style and different by-products, e.g., EyeSpy's by-product might instead be riddles.

Generalising from this, we suggest that it is difficult to guarantee that the by-products of this style of game will always be 'pure'. Given subjective tagging and confirmation, players may always potentially find new ways to play, or ways to 'game the system'. Designers can reduce the likelihood of such events but cannot avoid them completely. Careful design and thorough testing should help, but we recommend that designers also consider the ways in which the language used to introduce the game and frame the system will influence the production of particular kinds of by-products.

### **Human 'algorithms'**

This relationship also raises more questions about the growing body of literature documenting the design and implementation of systems exploiting so-called 'human computation' or 'human algorithms' [21]. EyeSpy may be seen as a solution to a machine vision problem of selecting relevant images for navigational tasks within a geographical location. However, we would argue that it also reveals a number of challenges for designers creating systems within this domain.

Firstly, we should be careful about not taking phrases such as humans as "processing nodes for problems" [21] too literally. While we agree that there are great opportunities for human solutions to hard computational problems, there are also considerable design challenges which must be addressed when combining human and machine 'computation'. In particular, player motivation is a key component to such systems' success but this can be lost easily. People may get bored by the simplicity of a game, or confused by over-complex rules. They may suffer from fatigue, or have problems weaving their play into everyday life. Other significant social factors will influence the success of human computation systems, such as how to 'market' them to potential users.

The *accountability* of these systems also plays a key design role. Algorithms are generally deterministic and have known upper bounds calculation time. They are highly 'accountable' in that one can examine in detail precisely how an output was created. In comparison, within games such as EyeSpy and the ESP game, for example, the time needed to obtain information is subject to the vagaries of player participation, motivation and conformity with regard to norms of play. However, in contrast to the fixed accountability of an algorithm, this form of accountability is negotiated continuously between players themselves. Designers should

be aware of the possibilities of both gaming the system and feeding the system with spurious data (as mentioned by others [20]). Apart from design, testing and 'framing the game', such problems highlight the importance of moderation, quality control and 'orchestration' activities as vital components in keeping the system running successfully [10].

Finally, designers should be aware of the way in which specificities of individuals and groups may impact how human computation systems work out in practice. Within EyeSpy, exploiting 'what anyone knows' involved drawing on local knowledge in order to successfully capture navigationally useful images. This was exposed particularly well by one of the participants in his orientation to 'students' as hypothesised recipients of his images. Conflicts in 'what anyone knows' can also come to bear when categorising the content of images, for instance; we can imagine how one symbol may mean very different things to different groups of users (e.g., a swastika or manji, commonly used in Japanese maps to mark temples). Human computation then is not just about producing 'objective' results, but can also be about using subjective understandings to produce content that draws upon subjective, creative knowledges. In EyeSpy this was simply judgements on what people could find in a local area – but even this depends on the cultural positioning of players (e.g. as pedestrians in the city rather than drivers). In EyeSpy exploiting the local knowledge of participants simply meant producing more culturally relevant images; the 'algorithm' employed in the game was in a sense truly adaptive to its context of use.

### **FUTURE WORK**

Following on from EyeSpy, we are investigating ways to address issues such as sociability and saturation. The most obvious way to improve sociability is to make it a team game. For example, teams may find ways to coordinate and combine their tagging, and enjoy the social interaction of collaboration. One approach to dealing with the saturation of popular areas that we are considering is to have tags fade over time, so that players will have to revisit (and thereby keep up-to-date) the tags in areas they find convenient to play in. Alternatively, we could increase the pay-off for players who explore or 'open up' new areas for play. Given the potential variability of where players may go and what they may consider to be 'good tags' within a long-term game, we are also investigating ways to dynamically 'steer' the scoring scheme and information about other players. For example, we may mark areas that have not been tagged enough as being worth extra points, or we may hide existing tags in a given area so as to convince new players to play there. A difficult issue here is how to keep the game balanced despite changes to its structure. For example, a player who has saturated an area and so built up a lead in the game may object to new and distant areas opening up that may let others win.

An iPhone version of EyeSpy is also in development and we plan to make this version freely available to encourage as wide a range of users as possible. Potentially, if widely used, the game could automatically collect navigation images worldwide, thus producing a valuable corpus through enjoyable end user experiences, as well as enabling us to test EyeSpy on a wider scale and in a greater variety of contexts.

## CONCLUSION

We have explored issues surrounding the design, implementation and testing of a simple pervasive game that produces by-products that may be repurposed for tasks involving navigation. By bringing verification into the mobile game, we were able to produce one of the first full examples of human computation in a mobile context. The technical implementation of the game took advantage of the availability and speed of wifi in order to geo-locate players and their tags. Our trial of the game demonstrated players' orientation to certain navigational qualities—i.e., recognisability and findability—of the tags they created, as well as revealing the careful design balance between game rules, how rules work out in practice, and the character of the by-products that are produced. We carried out two experiments assessing the general character of the image set generated by the game, and the practical use of the images in a simple navigational task. They confirmed that EyeSpy did indeed produce images that were more recognisable and findable than a popular archive of geo-located images. Finally we reflected upon the design challenges and opportunities posed by human computation, drawing attention to the delicate balance between rules and their by-products, as well as human issues of designing for fun, motivation, maintaining interest, accountability, and cultural specificities. We consider human computation to have great promise, both in mobile and more traditional networked settings, and games to be particularly promising vehicles for it, even though further experience will be needed before this burgeoning design paradigm can meet its full potential.

## ACKNOWLEDGEMENTS

This research was funded by UK EPSRC (EP/F035586/1, EP/E04848X/1, GR/N15986/01). Thanks also to the other members of the Social/Ubiquitous/Mobile Group who helped directly or indirectly with this work.

## REFERENCES

1. Barkhuus, L. et al. Picking pockets on the lawn: the development of tactics and strategies in a mobile game. In *Proc. Ubicomp*, pp. 358-374, 2005, Springer.
2. Beeharee, A. and Steed, A. Minimising Pedestrian Navigational Ambiguities Through Geoannotation and Temporal Tagging. In *Human-Computer Interaction. Interaction Platforms and Techniques*, Springer, 2007, 748-757.
3. Bell, M. et al. Interweaving mobile games with everyday life. In *Proc. ACM CHI*, pp. 417-426, 2006.
4. Benford, S. et al. Can you see me now? *ACM ToCHI*: 13(1), pp. 100-133, 2006.
5. Benford, S. et al. 2000. Designing storytelling technologies to encourage collaboration between young children. In *Proc. ACM CHI*, pp. 556-563, 2000.
6. Brown, B. and E. Laurier. Designing electronic maps: an ethnographic approach. L. Meng, A. Zipf, T. Reichenbacher (Eds.), *Map-based mobile services – Theories, Methods and Implementations*, pp. 247-265, Springer, 2005.
7. Garfinkel H. Some rules of correct decision making that jurors respect. In *Studies in Ethnomethodology*, pp. 104-115, 1967, Prentice Hall.
8. Goldin, D. et al. *Interactive Computation: The New Paradigm*. Springer Verlag: 2006.
9. Grant, L. et al. MobiMissions: the game of missions for mobile phones. In *ACM SIGGRAPH Educators Program*, 2007.
10. Koleva, B. et al. Orchestrating a mixed reality performance. In *Proc. ACM CHI*, pp. 38-45, 2001.
11. Leshed, G. et al. In-car GPS navigation: engagement with and disengagement from the environment. In *Proc. ACM CHI*, pp. 1675-1684, 2008.
12. Laurier, E. and Brown, B. Rotating maps and users: praxiological aspects of alignment and orientation. *Trans. Inst. British Geographers*, 33 (201-221), 2006.
13. Matyas, S. Playful Geospatial Data Acquisition by Location-Based Gaming Communities. *IJVR*. IPI Press: 6(3), pp. 1-10, 2007.
14. Matyas, S. et al. Designing Location-based Mobile Games with a Purpose—Collecting Geospatial Data with CityExplorer. In *Proc. ACE*, 2008.
15. Mor, N. et al. Photos on the go: a mobile application case study. In *Proc. ACM CHI*, pp. 1739-1748, 2008.
16. O'Hara, K. Understanding geocaching practices and motivations. In *Proc. ACM CHI*, pp. 1177-1186, 2008.
17. Petra, J. et al. In the hands of children: exploring the use of mobile phone functionality in casual play settings. In *Proc. Mobile HCI*, 2008.
18. Sacks, H. et al. A simplest systematics for the organization of turn taking for conversation. In *Studies in the organization of conversational interaction*, pp. 7-55, 1978, Academic Press.
19. Tollmar, K. et al. IDEixis – Searching the Web with Mobile Images for Location-Based Information. In *Proc. Mobile HCI*, pp. 288-299, 2004.
20. von Ahn, L. and Dabbish, L. Labeling images with a computer game. In *Proc ACM CHI*, pp. 319-326, 2004.
21. von Ahn, L. et al. Verbosity: a game for collecting common-sense facts. *Proc. ACM CHI*, pp. 75-78, 2006.