Psychological Maps 2.0: A Web Engagement Enterprise Starting in London

Daniele Quercia
Yahoo! Research, Barcelona
dquercia@yahoo-inc.com

João Paulo Pesce
UFMG, Brazil
jpesce@dcc.ufmg.br

Virgilio Almeida
UFMG, Brazil
virgilio@dcc.ufmg.br

Jon Crowcroft
University of Cambridge, UK
jon.crowcroft@cl.cam.ac.uk

ABSTRACT

Planners and social psychologists have suggested that the recognizability of the urban environment is linked to people's socio-economic well-being. We build a web game that puts the recognizability of London's streets to the test. It follows as closely as possible one experiment done by Stanley Milgram in 1972. The game picks up random locations from Google Street View and tests users to see if they can judge the location in terms of closest subway station, borough, or region. Each participant dedicates only few minutes to the task (as opposed to 90 minutes in Milgram's). We collect data from 2,255 participants (one order of magnitude a larger sample) and build a recognizability map of London based on their responses. We find that some boroughs have little cognitive representation; that recognizability of an area is explained partly by its exposure to Flickr and Foursquare users and mostly by its exposure to subway passengers; and that areas with low recognizability do not fare any worse on the *economic* indicators of income, education, and employment, but they do significantly suffer from *social* problems of housing deprivation, poor living conditions, and crime. These results could not have been produced without analyzing life off- and online: that is, without considering the interactions between urban places in the physical world and their virtual presence on platforms such as Flickr and Foursquare. This line of work is at the crossroad of two emerging themes in computing research - a crossroad where "web science" meets the "smart city" agenda.

Categories and Subject Descriptors

H.4 [**Information Systems Applications**]: Miscellaneous

Keywords

Social Media, Web Science, Urban Informatics

1. INTRODUCTION

A geographic map of a city consists of, say, streets and buildings and reflects an *objective* representation of the city. By contrast, a psychological map is a *subjective* representation that each city dweller carries around in his/her head. Tourists in a strange city start with few reference points (e.g., hotels, main streets) and then expand the representation in their minds - they slowly begin to build a picture. To see how these subjective representation matter, consider London. Every Londoner has had long attachment with some parts of the city, which brings to mind a flood of associations. Over the years, London has been built and maintained in a way that it is imaginable, i.e., that mental maps of the city are clear and economical of mental effort. That is because, starting from Kevin Lynch's seminal book "The Image of the City" in 1960, studies have posited that good imaginability allows city dwellers to feel at home and increase their community well-being [13]. People generally feel at home in cities whose neighborhoods are recognizable. Comfort resulting from little effort, the argument goes, would impact individual and ultimately collective well-being.

The good news is that the concept of imaginability is *quantifiable*, and it is so using psychological maps (Section 2). Since Stanley Milgram's work in New York and Paris [17, 16], researchers (including HCI ones) have drawn recognizability maps by recruiting city dwellers, showing them scenes of their city, and testing whether they could recognize where those scenes were: depending on which places are correctly recognized, one could draw a collective psychological map of the city. The problem is that such an experiment takes time (in Milgram's, each participant spent 90 minutes for the recognition task), is costly (because of paid participants), and cannot be conducted at scale (so far the largest one had 200 participants). That is why the link between recognizability of a place and well-being of its residents has been hypothesized, qualitatively shown, but has never been quantitatively tested at scale.

To test whether the recognizability of a place makes it a more desirable part of the city to live, we make the following contributions:

First, we build a crowdsourcing web game that puts the recognizability of London's streets to the test (Section 3). It picks up random locations from Google Street View and tests users to see if they can determine in which subway location (or borough or region) the scene is. In the last five

Preprint. To appear in ACM WWW 2013.

months, we have collected data from 2,255 users, have built a collective recognizability map of London based on their responses, and quantified the recognizability of different parts of the city.

Second, by analyzing the recognizability of London regions (Section 4), we find that the general conclusions drawn by Milgram for New York hold for London with impressive consistency, suggesting external validity of our results. Central London is the most recognizable region, while South London has little cognitive coverage. Londoners would answer "West London" when unsure, making the most incorrect guesses for that region - hence a West London response bias. We also find that the mental map of London changes depending on where respondents are from - London, UK, or rest of the world.

Third, we test to which extent an area's recognizability is explained by the area's exposure to people (Section 5). In particular, we study exposure to users of three social media services and to underground passengers. By collecting 1.2M Twitter messages, 224K Foursquare check-ins, 76.6M underground trips, and 1.3M Flickr pictures in London, we find that, the more a social media platform's content is geographically salient (e.g., Flickr's), the better proxy it offers for recognizability.

Finally, upon census data showing the extent to which areas are socially deprived or not, we test whether recognizability of an area is negatively related with the area's socio-economic deprivation (Section 6). We find that recognizability is indeed low in areas that suffer from housing deprivation, poor living conditions, and crime.

This work is at the crossroad of two emerging themes in computing research - web science and smart cities. The combination of the two opens up notable opportunities for future research (Section 7).

2. BACKGROUND

Psychological maps from drawing one's version of the city. In his 1960 "The Image of the City", Kevin Lynch created a psychological map of Boston by interviewing Bostonians. Based on hand-drawn maps of what participants' "versions of Boston" looked like, he found that few central areas were know to almost all Bostonians, while vast parts of the city were unknown to its dwellers. More than ten years later, Stanley Milgram repeated the same experiment and did so in a variety of other cities (e.g., Paris, New York). Milgram was an American social psychologist who conducted various studies, including a controversial study on obedience to authority and the original small world (six degree of separation) experiment [15]. Milgram was interested in understanding mental models of the city, and he turned to Paris to study them: his participants drew maps of what "their versions of Paris" looked like, and these maps were combined to identify the intelligible and recognizable parts of the city. Since then, researchers have collected people's opinions about neighborhoods in the form of hand-drawn maps in different cities, including (more recently) San Francisco [1] and Chicago [3].

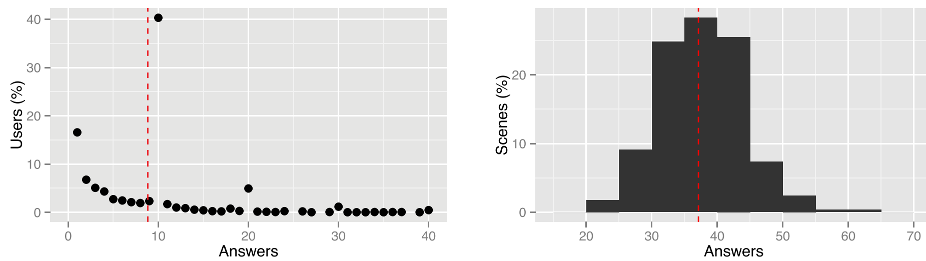
Psychological maps from recognizing city scenes. The problem with the mental map experiment is that it takes time and it is not clear how to aggregate the variety of unique configurations of answers that are bound to appear. One way of fixing that problem is to place a number of constraints on the participants when externalizing their maps. In this vein, before his experience with Paris, Milgram constrained the experiment so much so to reduce it to a simple question: "If an individual is placed at random at a point in the city, how likely is he to know where he is?" [17]. The idea is that one can measure the relative "imaginability" of cities by finding the proportion of residents who recognize sampled geographic points. That simply translates into showing participants scenes of their city and testing whether they can recognize where the scenes are. Milgram did setup and successfully run such an experiment in lecture theaters. Each participant usually spent 90 minutes on the task, and he collected responses from as many as 200 participants for New York. Hitherto the experimental setup in which maps are drawn has been widely replicated [1, 3], while that in which scenes are recognized has received far less attention. Next, we try to re-create the latter experimental setup at scale by building an online crowdsourcing platform in which each participant plays a one-minute game, a game with a purpose [23].

Figure 1: Google Street View Scene.

3. PSYCHOLOGICAL MAPS 2.0

We have created an online game that asks users to identify Google Street View (Panorama) scenes of London. The project aims to learn how its players mentally map different locations around the city, ultimately creating a London-wide map of recognizability.

How it works. For each round, the game shows a player a randomly-selected scene in London and ask him/her to guess the nearest subway station, or generally what section of the city (borough/region) (s)he is seeing. Answers should be easy, and that is why we choose the finest-grained answer to be subway stations as those are the most widely-used point of references among Londoners and visitors alike. To avoid sparsity problems (too few answers per picture), a random scene is selected within a 300-meter radius from a tube station but not next to it (to avoid easing recognizability). The idea is that, by collecting a large number of responses across a large number of participants, we can determine which areas are recognizable. By testing which places are remarkable and unmistakable and which places represent faceless sprawl, we are able to draw the recognizability map of London.



(a) Number of Answers per User (b) Number of Answers per Scene

Figure 2: Number of Answers for Each User/Scene. *(a)* Each game round consists of 10 pictures - that is why outliers are multiples of 10. The analysis considers the 40% of users who have completed one round. *(b)* The number of answers each scene has received is normally distributed thanks to randomization.

Engagement strategies. The strategies we implemented include:

Giving Points. "One of the most direct methods for motivating players in games is to grant points for each instance of successful output produced. Using points increases motivation by providing a clear connection among effort in the game, performance, and outcomes" [23]. When playing the game, each player receives a score that increases with the number of correct guesses of where a given picture was taken. To enhance the gaming experience, we reward not only strictly correct guesses (which are the only ones considered for experimental sake) but also "geographically close" ones by awarding points based on the Euclidian distance d between a user's guess and the correct answer. The idea is that guesses within a radius of 300 meters still amount to reasonable scores, while those outside it are severely and increasingly penalized depending on how far they are from the correct answer. To reduce the number of random guesses, we allow for an "I Don't Know" answer, which still rewards players with 15 points. After being presented with 10 pictures, the player has completed one round and (s)he can share the resulting score on Facebook or Twitter with only one click. The score is supposed to facilitate the player's assessment of his/her performance against previous game rounds or against other players [23]. From the distribution of number answers for each player (Figure 2(a)), we find multiples of 10 to be outliers, suggesting that players do tend to complete at least one round. After the first round, each player is also shown a small questionnaire (e.g., age, gender, location) (s)he is asked to complete. Participants engaging in multiple rounds are identified through browser cookies, which uniquely identify users [1].

Social Media Integration. Players can post their scores on the two social media platforms of Facebook and Twitter after each round with a default message of the form "How well do you know London? My score . . . ". The goal of such a message is to make Facebook and Twitter users aware of the game.

Randomness. "Games with a purpose should incorporate randomness. For example, inputs for a particular game session are typically selected at random from the set of all possible inputs. Because inputs are randomly selected, their difficulty varies, thus keeping the game interesting and engaging for expert and novice players alike." [23]. In our game, pictures are chosen randomly with the hope of creating a sense of freshness and increasing replay value. In addition, randomizing the selection of picture is a good idea for experimental sake. Randomization reduces spatial biases and leads to reliable results, producing a distribution of answers for each picture that is not skewed. In our analysis, such a distribution will turn out to be distributed around a mean of 37.11 (Figure 2(b)), and no picture has less than 20 answers.

Overall, by providing a clear sense of progression and goals that are challenging enough to maintain interest but not so hard as to put players off, we hope to capture a sense of engagement.

From beta to final version. We build a working prototype featuring those desirable engagement properties and ascertain the extent to which it works in a controlled beta test involving more than 45 urban planners, architects, and computer scientists. We receive four main feedbacks:

Ease. The game is found to be difficult and, as such, frustrating to play. That is because random pictures from every (remote) part of London are shown. One player said: "I've been living in London for the past 35 years and I felt like a tourist. There were so many places I had no clue where they were. It is frustrating to get a score of 200 [out of 1000]! ". To fix this problem, we manually add pictures of easily recognizable places (e.g., spots that are touristic or close to subway stations) and show them together with the randomly selected places from time to time. For the purpose of study, these "fake" pictures are ignored - they are just meant to improve the gaming experience and retention rate.

Feedbacks. The beta version does not show any feedback about which are the correct answers. A large number of testers feel that the game could be an opportunity to learn more about London. That is why, for each incorrect guess, the final version of the game also shows the right answer.

Sense of purpose. The site does not contain any explanation about the research aims behind the game. Yet, our testers feel that providing a sense of purpose to players was essential. The final version contains a short explanation of how the game is designed for purposes beyond pure entertainment, and how it might be used to promote urban interventions where needed.

[1] Unless two users use the same computer and the same username on it. This situation should represent a minority of cases though.

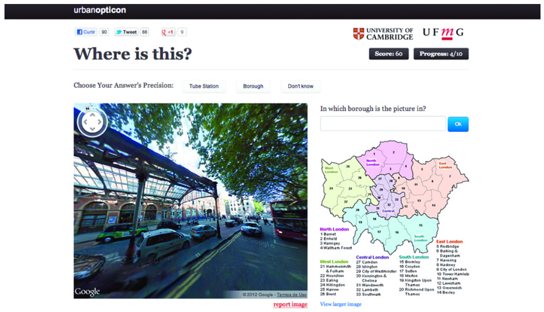


Figure 3: Screenshot of the Crowdsourcing Game. The city scene is on the left, and the answer box on the right.

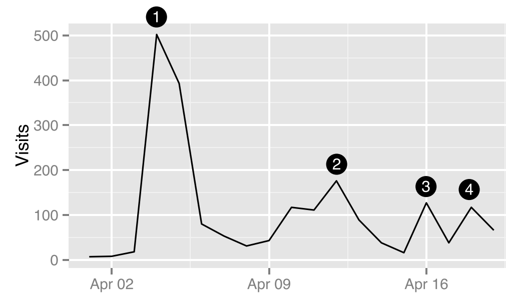


Figure 4: Initial Visitors on the Site. Peaks are registered for: (1) Cambridge press release; (2) The Independent article; (3) New Scientist article; and (4) residual sharing activity on Facebook and Twitter.

	London	UK	World	Total
Answers	7,238	8,705	3,972	19,915
Users	739	973	543	2,255
Gender (%)				
Male	59.1	64.3	46.5	59.6
Female	40.9	35.7	53.5	40.4
Age (%)				
<18	0.9	0.8	0.0	0.7
18-24	16.5	24.8	9.3	19.2
25-34	41.8	38.9	51.2	41.8
35-44	16.5	13.9	20.9	16.0
45-54	13.9	13.9	7.0	12.9
55-64	5.2	6.2	9.3	6.3
65+	5.2	1.5	2.3	3.1
Mean (years)	36.4	33.9	34.5	35.0

Table 1: Statistics of Participants. Gender and age are available for those 287 participants (13%) who have been willing to provide personal details.

Beyond one type of answer. The game asks players to guess the correct subway station. Many testers feel the need for coarser-grained answers. “I know this is Westminster [a borough in London], but I have no idea of the exact tube station!”, says one player. The final version thus allows for multiple types of answers: not only subway stations but also boroughs (50 points) or regions such as Central London and South London (25 points).

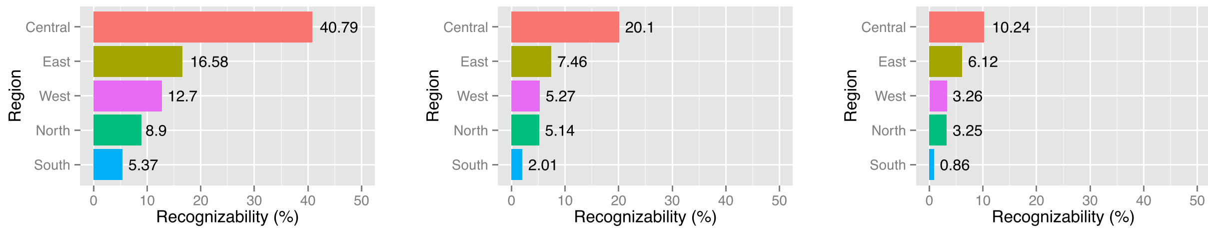
To sum up, the final version of the game works by giving a player ten (random plus morale boosting) images in Greater London (Figure 3). The player can either guess the tube station, borough, region, or click “Don't know” to move ahead. At the end of the round, the player is given a total score based on the fraction of correct answers. The score can be automatically shared on Facebook or Twitter, and the player is presented with a survey that asks for personal details like birth location, place of employment, and familiarity with the city itself.

Launching the crowdsourcing game. We have made the final version of the game publicly available and have issued a press release in April 2012 (Figure 4). Shortly after that, the game has been featured in major newspapers, including The Independent (UK national newspaper) and New Scientist. After 5 months, we have collected data from as many as 2,255 participants: 739 connecting from London (IP addresses), 973 from the rest of UK, and 543 outside UK. A fraction of those participants (287) specified their personal details. The percentage of male-female participants overall is 60%-40% and slightly changes depending on one's location: it stays 60%-40% in London but changes to 65%-35% in UK and 45%-55% outside it. Also, across locations, average age does not differ from London's, which is 36.4 years old. As for geographic distribution of respondents, we find a strong correlation between London population and number of respondents across regions ($r = 0.82$). Having this data at hand, we are now ready to analyze it.

4. RELATIVE RECOGNIZABILITY

The goal of this project is to *quantify* the relative recognizability of different parts of London. Since familiarity with different parts of the city might depend on place of residence, we filter away participants outside London and consider Londoners first. According to the Greater London Authority's division, London is divided into five different city (sub) regions. Thus, our first research question is to determine which proportion of the Google Street View scenes from each region were correctly attributed to the region. Since users were asked to name either the borough of each scene or the subway station closest to it, we consider an answer to be correct, if the region of the scene is the same as the region of the answered borough/subway station. For each of the five regions, we compute the region's percentage *recognizability* by summing the number of correct answers and then dividing by the total number of answers. Figure 5 reports the results. Clearly, Central London emerges as the most recognizable of the five regions, with about two and a half as many correct placements as the others. Conventional wisdom holds that Central London is better known than other parts of the city, as it hosts the main squares, major railway and subway stations, and most popular touristic attractions and night-life “hotspots”. Interestingly, the East Region is twice as recognizable than the North Region. It is difficult to draw conclusions on why this is. However, the three most likely explanations are:

Sample Bias. It might depend on the distribution of the home and work addresses of our participants. However, that



(a) *Region*-level recognizability (b) *Borough*-level recognizability (c) *Subway station*-level

Figure 5: Recognizability Across London Regions. This is computed based on whether scenes are recognized at: *(a) region* level; *(b) borough* level; or *(c) subway station* level.

is unlikely, as the correlation between London population across regions and number of participants who answered the survey is as high as $r = 0.82$. Despite that correlation, we cannot fully rule out the sampling bias though.

Large volume of visitors. High recognizability for the East part of the city can be explained by an experiential effect. Large numbers of people are expected to visit that part of the city: workers at Canary Wharf, visitors to Olympic Park, Excel, City airport, and O2 arena. A recent study of Londoners' whereabouts on Foursquare found them to be skewed towards mostly Central London and partly East London - especially the central east part [2]. The north parts are unlikely to have been visited by similar volumes of people. In Section 5, we will see that there is a significant correlation between recognizability of an area and the area's exposure to specific subgroups of individuals. For example, we will see that the more passengers use an area's subway station, the more recognizable the area ($r = 0.45$).

Distinctiveness of the built environment. The East region includes most of the City and Canary Wharf (financial area with skyscrapers), as well as the O2 arena and Docklands region, all more visually recognizable areas than comparable parts of the North region. Also, East London has been affected by large homogeneous post-war housing projects that make the area quite distinctive [9].

Next, we adopt a more stringent criterion of recognition, that is, we determine what proportion of the scenes in each of the five regions were placed in the correct *borough.* By analyzing the answers at borough-level, we find substantial differences (Figure 5(b)). A scene placed in Central London is almost three times more likely to be placed in the correct borough than a scene in East London, and are four times more likely than a scene in West or North London.

When we then apply even a more stringent criterion of recognition (subway station), the correct guesses are drastically reduced (Figure 5(c)), as one expects. Interestingly, the information value of Central London is less pronounced. Central London scenes are only one and a half time more likely to be associated with the correct subway station than a scene in East London. Guessing the correct subway stations is hard, the more so in the central part of the city where stations are close to each other. During post-game interviews, one participants noted: "Perhaps people know where places are, but have difficulty identifying which of the [subway stations] it is actually close to." Despite these differences, the *relative* recognizability (ranked recognizability of the five regions) does not change. Figure 6 shows the

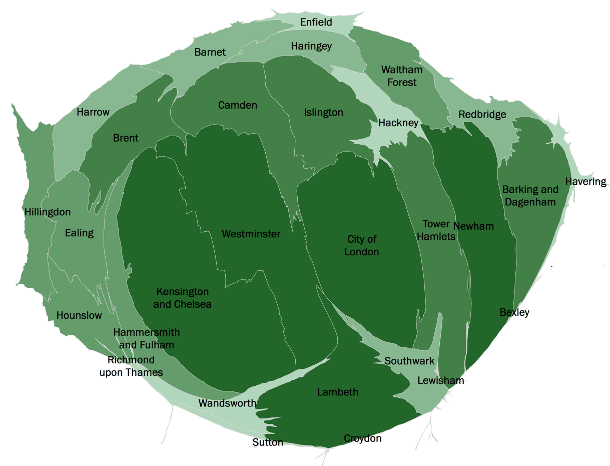


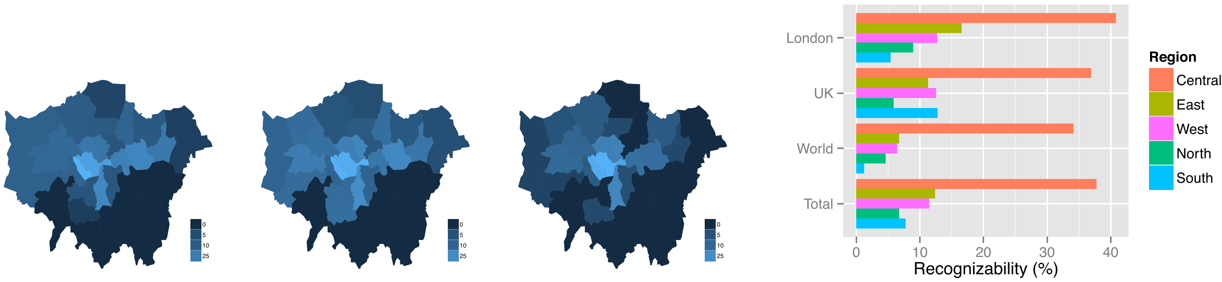
Figure 6: Cartogram of London Boroughs. The geographic area is distorted based on borough's recognizability.

cartogram of London boroughs. The geometry of the map is distorted based on recognizability scores. Central London dominates, while South London is relegated at the bottom.

Another aspect to consider is that one is likely to recognize areas closer to where one lives or works. Based on our survey respondents, we find that there is no relationship between recognizability of a scene and a respondent's self-reported home location. On the contrary, participants are more likely to recognize scenes in Central London rather than scenes in their own boroughs.

The recognizability of each region does not change depending on which parts of the city Londoners live, but does change depending on whether participants are in UK or not. Based on our participants' IP addresses [2], we infer the cities where they are connecting from, and compute aggregate correct guesses by respondent location - that is, by whether participants connect from London, from the rest of UK, or outside UK (Figure 7). As expected, the number of correct guesses drastically decreases for participants outside London - but with two exceptions. First, scenes of South London are more recognizable for participants in the rest of UK than for Londoners themselves. That is because Londoners tend

[2]There might be cases of misclassification of cities and of people who use VPNs. However, at the three coarse-grained levels of London *vs.* rest of UK *vs.* rest of the world, misclassification should have a negligible effect.



(a) Londoners (b) UK (c) Outiside UK (d) All

Figure 7: Recognizability Across London Regions by Respondent Location. On the maps, lighter colors correspond to more recognizable places.

Region actually is	But identified as C	E	W	N	S	Combined Errors	Don't Know
Central	**40.79**	4.52	4.33	1.03	2.13	12.02	47.19
East	6.97	**16.58**	6.80	6.30	7.46	27.53	55.89
West	10.10	6.42	**12.70**	5.77	5.92	28.21	59.09
North	6.85	4.79	12.67	**8.90**	7.53	31.85	59.25
South	6.04	5.37	11.41	3.36	**5.37**	26.17	68.46
Response Bias*	29.96	21.1	**35.21**	16.46	23.04		

* popular among wrong guesses

Table 2: Matrix of Correct Classifications and Misclassifications.

to know Southfields (known as "The Grid", which a series of parallel roads that consist almost entirely of Edwardian terrace houses), while people in the rest of UK recognize scenes not only in Southfields, but also in Clapham South, Balham, South Wimbledon, and Tooting Broadway in that order. Second, recognizability of Central London remains the same across participants from all over: participants outside UK are as good as those inside it at recognizing scenes in Central London. Hosting the most popular tourist attractions in the world, Central London is vividly present in the world's collective psychological map.

So far we have focused on correct guesses. Now we turn to errors that respondent often make, looking for widely-shared sources of confusion. We wish to know in which regions (e.g., North, South) a scene from, say, East London is often misplaced. To this end, Table 2 shows a matrix reporting both the percentage of correct guesses and that of wrong ones for each region. Central London is pre-eminent in Londoners' shared psychological maps as it is hardly confused with any other region. At times, instead, South and North London are thought to be West. It seems that, if respondents do not know where to place a scene, they would preferentially opt for West London. Indeed, the West part of the city is the most popular answer for those who end up guessing wrongly (last row in Table 2). We found a *West London response bias*, as Milgram would put it [3].

Summary. Taken together, the results suggest two generalizable principles on why people recognize an area. They do so because they are exposed to it (Central London attracts dwellers from all over the city), and because the area offers a distinctive architecture (e.g., stadium, tower building) or cultural life (as the central part of East London notoriously does). Milgram found the very same two principles to hold for New York as well in 1972. So much so that Milgram hypothesized that the extent to which a scene will be recognized can be described by $R = f(C \cdot D)$, where R is recognition (our recognizability), C centrality of population flow (in the next section, we will see how to compute flow of subway passengers), and D is the social or architectural distinctiveness. It follows that, with simplifying assumptions (e.g., f is a linear relationship), one could derive an area's social or architectural distinctiveness by simply dividing recognizability by subway passenger flow. Since we are interested in the *relative* recognizability and flow, we take the rank values for these two quantities, compute their ratio, and report the results in Table 3. The most distinctive area is Blackfriars. It should be no coincidence that its older parts happen to "have regularly been used as a filming location in film and television, particularly for modern films and serials set in Victorian times, notably Sherlock Holmes and David Copperfield" [4]. In line with Milgram's experiment with New Yorkers [17], we find that the acquisition of a mental map is not necessarily a direct process but can also be indirect through, for example, movies. The following quote from one of our participants is telling: "I've done the quiz 3 or 4 times in the last couple of days, and am surprised how well I am doing - not just because I live in New Zealand, many thousands of kilometres from London (although I did live there for 10 years), but mainly because I am getting good scores on parts of London I have never been to. North and Eastern boroughs like Brent and Haringey seem to be recognisable, even though I have never knowingly gone there. Possibly some recognition from TV programs, or just - could it be - that there is something intrinsically North London about certain types of houses? "

5. RECOGNIZABILITY AND EXPOSURE

5.1 Digital Data for Exposure

The goal of the game is to quantify the recognizability of the different parts of the city. It has been shown that New Yorkers are able to recognize an area partly because they were exposed to it [17]. Thus, to quantify the extent

[3] Milgram found that New Yorkers would opt for answering "Queens" when unsure - hence he referred to a "Queens response bias" [17]

[4] `http://en.wikipedia.org/wiki/Blackfriars,_London`

name	R	C	r_R	r_C	D
Blackfriars	9.09	4583	30	2	15.00
Park Royal	20.00	13119	61	5	12.20
Pinner	10.00	13823	37	6	6.17
Royal Oak	10.00	16681	37	8	4.63
Westbourne Park	16.66	24593	54	13	4.15
Hornchurch	7.14	11988	16	4	4.00
Essex Road	5.55	2027	4	1	4.00
Oakwood	11.11	22321	41	11	3.73
Hillingdon	6.67	9482	11	3	3.67
Acton Town	40.00	33022	73	22	3.32

Table 3: Subway Stations of Socially/Architecturally Distinctive Areas. For each area, R is the recognizability, C is the flow centrality (number of unique subway passengers), r_R and r_C are the corresponding ranked values, and D is the normalized distinctiveness.

to which it is so in London, we measure the exposure that an area receives by computing the number of overall unique individuals who happen to be in the area. These individuals are of four subgroups: those who post Twitter messages while in the area, those who visit locations (e.g., restaurants, bars) and say so on Foursquare, those who take pictures of the area and post them on Flickr, and those who catch a train in the closest subway station. We are thus able to associate the recognizability of an area with the area's exposure to these four subgroups.

Twitter geo-enabled users. Our goal is to retrieve as large and unbiased a sample of geo-referenced tweets as possible. To do this, we use the public streamer API, which connects to a continuous feed of a random sample of all ever shared tweets, and crawl geo-referenced tweets within the bounding box of Greater London. During the period that goes from December 25th 2011 to January 12th 2012, we retrieve 1,238,339 geo-referenced tweets posted by 57,615 different users.

Foursquare users. *Gowalla*, *Facebook Places*, and *Foursquare* are popular mobile social-networking applications with which users share their whereabouts with friends. In this work, we consider the most used social-networking site in London - Foursquare [2]. Users can check-in to locations (e.g., restaurants) and share their whereabouts. We consider the geo-referenced tweets collected by Cheng *et al.* [4]. They collected Twitter updates (single tweets) that report Foursquare check-ins all over the world. We take the 224,533 check-ins that fall into Greater London. Those check-ins are posted by 8,735 users.

Flickr users. We collect photo metadata from `Flickr.com` using the site's public search API. To collect all publicly available geo-referenced pictures in the Greater London area, we divide the area into 30K cells, search for photos in each of them, and retrieve metadata (e.g., tags, number of comments, and annotations). The final dataset contains metadata for 1,319,545 London pictures geo-tagged by 37,928 users. This reflects a complete snapshot of all pictures in the city as of December 21st 2011.

Subway passengers. In 2003, the public transportation authority in London introduced an RFID-based technology, known as Oyster card, which replaced traditional paper-based magnetic stripe tickets. We obtain an anonymized dataset containing a record of every journey taken on the London rail network (including the London Underground) using an Oyster card in the whole month of March 2010. A record registers that a traveler did a trip from station a at time t_a, to station b at time t_b. In total, the dataset contains 76.6 million journeys made by 5.2 million users, and is available upon request from the public transportation authority.

Demographics of the individuals under study. Activity analyzed in this paper clearly relates only to certain social groups, and the exclusionary aspect of certain segments of the population should be acknowledged. It would be thus interesting to compare the demographics of the different types of individuals we are studying here. From a recent Ignite report on social media [10], global demographics of Foursquare and Twitter show a pronounced skew towards university educated 25-34 year old women (66% women for Foursquare and 61% for Twitter), while those of Flickr show a pronounced skew towards university educated 35-44 year old women (54% women). The demographics of subway passengers is by far the most representative but is also slightly skewed towards male with above-average income in the two age groups of 25-44 and 45-59 [21]. Instead, demographics of our London gamers show a skew towards 25-34 year old men (60% men). Thus, compared to social media users, our gamers reflect similar age groups but are more likely to be men. This demographic comparison should inform the interpretation of our results.

5.2 Recognizability and Exposure

After computing each area's exposure to people of four subgroups[5] (i.e., to users of the three main social media sites and to underground passengers), we are now ready to relate the area's exposure to its recognizability. We compute the Pearson's product-moment correlation between recognizability and exposure, for all four classes of individuals. Pearson's correlation $r \in [-1, 1]$ is a measure of the linear relationship between two variables. We expect that the more a given class of individuals is representative of the general population, the higher its correlation with recognizability. When computing the correlation, if necessary (e.g., because of skewness), variables undergo a logarithmic transformation. Figure 8 shows the relationship between recognizability and exposure to the four classes of individuals, with corresponding correlation coefficients (which are all significant at level $p < 0.001$). To put results into context, we should say that the exposure measures derived from the three social media sites all show very similar pair-wise correlations with exposure to subway passengers ($r \approx 0.60$), yet their correlations with recognizability show telling differences. Given that subway passengers are slightly more representative of the general population than social media users [21], it comes as no surprise that they show the highest correlation ($r = 0.40$). Both Flickr and Foursquare users are also associated with robust correlations ($r = 0.36$

[5]By area, we mean UK census area also known as Lower Super Output Area, which we will introduce in Section 6.

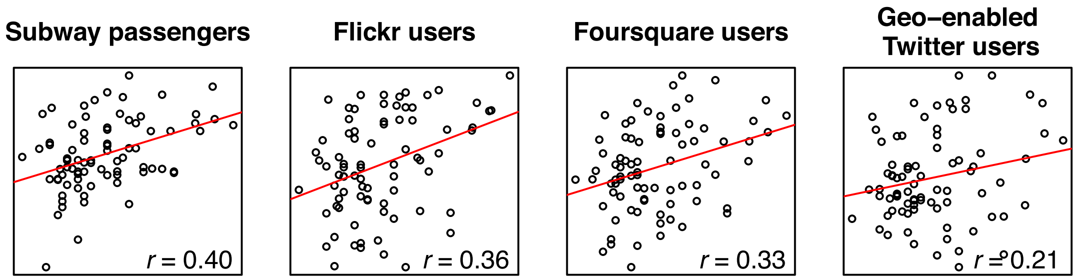


Figure 8: Area recognizability vs. Exposure to Four Classes of Individuals. Correlations are computed at *borough* level.

Exposure	Central	East	West	North	London
Subway	0.87	0.65	0.63	0.95	0.73
Flickr	0.62	0.50	0.27	0.89	0.63
Foursquare	0.72	0.36	0.22	0.97	0.58
Twitter	0.56	0.28	0.11	0.97	0.52

Table 4: Correlations between recognizability and Exposure by Region. Correlations are computed at *region* level. South London does not have enough subway stations to attain statistically significant correlations.

and $r = 0.33$). By contrast, having the least geographically salient content, Twitter shows a moderate correlation ($r = 0.21$). If we break the results down to regions (Table 4) and show which regions' recognizability is easy to predict from exposure and which not, we see that exposure to *any* subgroup of individuals would predict the recognizability of North London ($r = 0.95$). By contrast, the social media subgroup whose exposure correlates with recognizability the most in Central London is Foursquare ($r = 0.72$), and in East London is Flickr' ($r = 0.50$). That is largely because Foursquare activity is skewed towards Central London.

6. RECOGNIZABILITY AND WELL-BEING

As already mentioned in the introduction, Kevin Lynch outlined a theory connecting urban recognizability to a person's well-being [13]. To test this theory, we now gather census data on an area's socio-economic well-being and relate it to the area's recognizability.

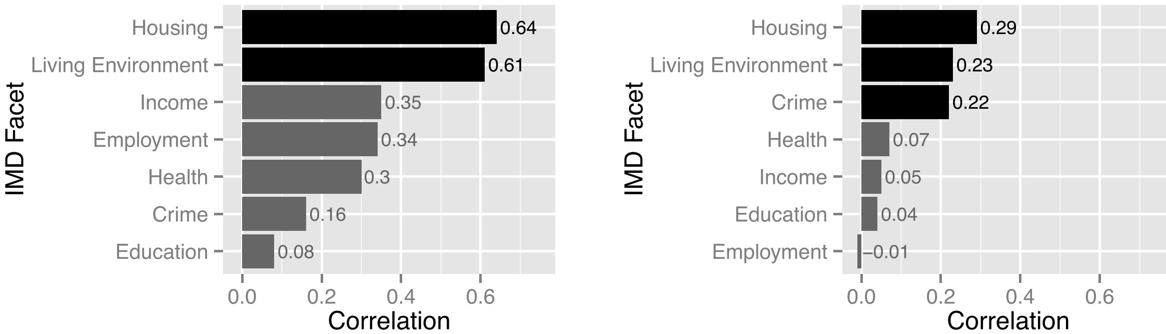
Facets of Socio-economic Well-being. Since 2000, the UK Office for National Statistics has published, every three or four years, the Indices of Multiple Deprivation (IMD), a set of indicators which measure deprivation of small census areas in England known as Lower-layer Super Output Areas [14]. These census areas were designed to have a roughly uniform population distribution so that a fine-grained *relative* comparison of different parts of England is possible. As per formulation of IMD, deprivation is defined in such a way that it captures the effects of several different factors. More specifically, IMD consists of seven components: 1. *Income* deprivation (e.g., number of people claiming income support, child tax credits or asylum); 2. *Employment* deprivation (e.g., number of claimants of jobseeker's allowance or incapacity benefit); 3. *Health* deprivation (e.g., including a standard measure of premature death, rate of adults suffering mood and anxiety disorders); 4. *Education* deprivation (e.g., education level attainment, proportion of working adults with no qualifications); 5. barriers to *Housing* and services (e.g., homelessness, overcrowding, distance to essential services); 6. *Crime* (e.g., rates of different kinds of criminal act); 7. *Living* Environment Deprivation (e.g., housing condition, air quality, rate of road traffic accidents); and finally a composite measure known as IMD which is the weighted mean of the seven domains.

Recognizability and Well-being. We start at borough level, correlate each transformed facet of deprivation [6] with recognizability, and obtain the results shown in Figure 9(a). We find that the composite score IMD does not correlate with recognizability at all. Neither does income, education, or (un)employment. What correlates are aspects less related to economic well-being and more related to social well-being: boroughs with low recognizability tend to suffer from housing deprivation ($r = 0.64$) and poor living environment ($r = 0.62$). Given the strong correlations [7], one could easily predict which boroughs suffer from housing deprivation and poor living conditions based on relative recognizability scores.

One might now wonder whether that would also be possible from social media data. We correlate each of the deprivation facets with exposure to the four subgroups (subway passengers plus users of three social media). For housing, we see that data on recognizability is hardly replaceable by social media data. Boroughs not suffering from housing deprivation (as per log-transformed score) are more recognizable ($r = 0.64$), and yet do not seem to be more exposed to our subgroups - all correlations between housing deprivation and exposure are not statistically significant. Instead, for living environment, we see that data on recognizability can be replaced by social media data. Boroughs with good living conditions are more recognizable ($r = 0.61$), and do tend to be more exposed to subway passengers ($r = 0.56$),

[6]To ease the interpretation of the correlation coefficients, we transformed (inverted) the deprivation scores, in that, the higher they are (e.g., high crime index), the better it is (e.g., low-crime area). We would thus expect the correlations between transformed deprivation scores and recognizability to be generally positive.

[7]Unless otherwise noted all correlations are significant at level $p < 0.001$.



(a) Borough Level

(b) Area Level

Figure 9: Correlation between recognizability and Deprivation at the Level of *(a)* Borough or *(b)* Area. For both, the composite index IMD is not shown as it does not correlate. Also, correlations significant at level $p < 0.001$ are shown with black (as opposed to grey) bars.

Flickr users ($r = 0.57$), Foursquare users ($r = 0.52$), and Twitter users ($r = 0.46,\ p < 0.01$).

Here we are not claiming that each census area in a borough is the same. If we were to say that, we would commit an ecological fallacy. For indicators that show high variability within a borough, however, there is a danger of committing such a fallacy. We therefore investigate correlations at the lower geographic level of census area. We correlate each facet of deprivation with recognizability and obtain the results shown in Figure 9(b). Again, the composite score IMD does not correlate with recognizability, while housing, living environment, and crime all do: areas with low recognizability tend to suffer from housing deprivation ($r = 0.29$), poor living environment ($r = 0.23$), and crime ($r = 0.22$). Crime has been added to the list of indicators associated with recognizability, and that is because crime is one of the deprivation facets that varies the most within a borough among the seven.

To sum up, from the previous results, we might say that, based on recognizability scores of *areas*, we could predict *whether* an area suffers from crime or not. Instead, based on recognizability scores of *boroughs*, one could predict not only *whether* but also *to which extent* a borough suffers from poor living conditions and housing deprivation. By contrast, social media data could only be used to identify boroughs with poor living conditions.

7. DISCUSSION

This work is deeply rooted in early urban studies but also taps into recent computing research, especially research on "games with a purpose", whereby one outsources certain activities (e.g., labeling images) to humans in an entertaining way [23]; research on large-scale urban dynamics [6, 7, 18]; and research on how location-based services affect people's behavior [3, 5, 12]. Initially, with this study, we were aiming at informing social media research in the urban context by establishing which social media data could be used as proxy for recognizability and exposure (key aspects in studies of urban dynamics). It turns out that the answer is complex, suggesting a word of caution on researchers not to take social media data at face value. However, there is one generalizable finding: the more the content is geographically salient (e.g., Foursquare's whereabouts *vs.* Twitter messages), the more it is fit for purpose.

7.1 Limitations

Control Variables. To increase response rate, we kept the survey as short as possible. It asks a minimum number of questions from which controlled variables are derived. However, this choice has drawbacks. For example, the survey asks for home location but does not ask for any other information about one's urban recognizability reach (the parts of the city one better knows visually). The problem is that one would know better (apart from the area one lives) also areas near work and on the way back home. We acknowledge this limitation but also stress that these differences are likely to cancel themselves out in a big sample like ours because of randomization.

Information Value of Scenes. Some pictures might be more revealing than others. The game has two kinds of pictures. The fake pictures (excluded from the analysis) are meant to increase retention rate and, as such, are easy to recognize - they depict touristic locations or well-known stations. The real pictures (included in the analysis) are instead less informative as they have been vetted by us. However, they might still contain clues that make them recognizable. We are currently discovering which visual cues tend to be associated with highly recognizable images (e.g., landmarks, memorable horrible buildings). We do so by automatically extracting image features, in a vein similar to an exploration recently proposed by Doersch *et al.* [8]. We are also discovering which visual cues tend to be associated with beautiful, quiet, and happy urban sceneries [19].

7.2 Smart Cities Meet Web Science

The share of the world's population living in cities has recently surpassed 50 percent. By 2025, we will see another 1.2 billion people living in cities. The world is in the midst of an immense population shift from rural areas to cities, not least because urbanization is powered by the potential for enormous economic benefits. Those benefits will be only realized, however, if we are able to manage the increased complexity that comes with larger cities. The 'smart city' agenda is about the use of technological advances in physical and computing infrastructure to manage that complexity and create better cities. We will now discuss the ways in which this work suggests that the future of web scientists is charged with great potentials.

Planning urban interventions. We have shown that the relationship between recognizability and specific aspects of socio-economic deprivation is strong enough to identify boroughs suffering from high housing deprivation and poor living conditions, and also areas affected by crime. There is strong demand for making cities smarter, and the ability to identify areas in need could provide real-time information to, for example, local authorities. They could receive early warnings and identify areas of high deprivation quickly and at little cost, which is beneficial for cash-strapped city councils when planning renewal initiatives. However, before making any policy recommendation, recognizability data (based on a convenience sample) needs to be supplemented by other types of data - for example, by underground data [11, 20].

Making experiments on the web. By turning the execution of the experiment into a game, we have applied principles from games to a serious task and and have been consequently able to harness thousands of human brains. This might be fascinating to social science researchers, who must usually pay people to participate in their experiments. The game we have presented inverts that rule: players will happily fork out time for the privilege of being allowed to test their knowledge of London. Indeed, participants were rewarded with being able to test how well they knew London. One participant added: "Yesterday we had few friends over for dinner. I started to play the game on my laptop, and that escalated into a ridiculous competition among all of us that left my husband - the only Londoner in the room - quite injured, so to speak".

Rewarding schemes. We should design and test alternative engagement strategies. For now, we have focused on intrinsic (as opposed to extrinsic) rewards [24]. That is because recent psychological experiments (summarized in Werbach's latest book "For the Win" [24]) have suggested that "intrinsic rewards (the enjoyment of a task for its own sake) are the best motivators, whereas extrinsic rewards, such as badges, levels, points or even in some circumstances money, can be counter-productive" [22]. In this vein, it might be beneficial to build a similar game on crowdsourcing platforms where participants are paid (e.g., on Mechanical Turk) and test how different reward schemes affect the externalization of the mental map. Finally, more research has to go into determining which incentives make engagement sustainable.

Beyond London. Comforted by the encouraging results, we are starting to bring the game to other cities in the world. With their recent "smart city" initiatives, Rio de Janeiro and São Paulo are fit for purpose and are thus next on the list. At the moment, when rolling the platform out, the coverage of Google street view is required, and, being not automatic, two main aspects need to be customized: 1) selection of the geographic landmarks users need to recognize - subway stations might not work equally well in all cities; and 2) selection of seed easy-to-guess locations to avoid player frustration - one could make this step automatic by selecting popular city locations from, e.g., Wikipedia. We are currently working on an open-source platform in which those aspects are made automatic. In the short term, to encourage other researchers to join in and allow for research reproducibility, we make the aggregate statistics and the platform's source code publicly available [8].

8. CONCLUSION

In the sixties, scholars started to design experiments that captured the psychological representations that dwellers had of their cities. In mid-2012, we have translated their experimental setup into a 1-minute web game with a purpose, and have began with a deployment in London. We have gained insights into the differing perceptions of London that are held by not only Londoners but also people in UK and the rest of the world. The pre-eminence of Central London in the world's collective psychological map speaks to the popularity of its landmarks and touristic locations. The acquisition of a mental map is a slow process that does not necessarily come from direct experience but might be indirectly learned from, for example, atlases or movies. It comes as no surprise that Blackfriars, having being often used as a filming location, turned out to be the most socially/architecturally distinctive area - that is, an area whose recognizability is explained less by exposure to people and more by its distinctiveness. We have been able to quantitatively show the extent to which Londoners' collective psychological map tallies with the socio-economic indicators of housing deprivation, living environment conditions, and crime. By then comparing different social media platforms, we have suggested that a platform's demographics and geographic saliency determine whether its content is fit for urban studies similar to ours or not. This is a preliminary yet useful guideline for the web community who has recently turned to the study of large-scale urban dynamics derived from social media data. In the long term, having our design suggestions and source code at hand, researchers around the world might well seize the opportunity to take psychological maps to other cities.

Acknowledgement. This research has been guided from nonsense to sense by a large number of individuals. We are eager to spread some credits (while of course remaining accountable for all blame) to Mike Batty, Henriette Cramer, Tim Harris, Tom Kirk, Daniel Lewis, Ed Manley, Amin Mantrach, Guy Marriage, Yiorgios Papamanousakis, Alan Penn, Kerstin Sailer, Chris Smith, and the members of the two mailing lists Space Syntax and Cambridge NetOS. This research was partially supported by RCUK through the Horizon Digital Economy Research grant (EP/G065802/1) and by the EU SocialSensor FP7 project (contract no. 287975).

9. REFERENCES

[1] R. Annechino and Y.-S. Cheng. Visualizing Mental Maps of San Francisco. *School of Information, UC Berkeley*, 2011.

[2] A. Bawa-Cavia. Sensing The Urban: Using location-based social network data in urban analysis. In *Pervasive Urban Applications (PURBA)*, 2011.

[3] F. Bentley, H. Cramer, W. Hamilton, and S. Basapur. Drawing the city: differing perceptions of the urban

[8] http://profzero.org/urbanopticon/

environment. In *Proceedings of ACM Conference on Human Factors in Computing Systems (CHI)*, 2012.

[4] Z. Cheng, J. Caverlee, and K. Lee. Exploring millions of footprints in location sharing services. In *Proceedings of the AAAI International Conference on Webblogs and Social Media (ICWSM)*, 2011.

[5] H. Cramer, M. Rost, and L. E. Holmquist. Performing a check-in: emerging practices, norms and 'conflicts' in location-sharing using foursquare. In *Proceedings of ACM International Conference on Human Computer Interaction with Mobile Devices and Services (MobileHCI)*, 2011.

[6] D. J. Crandall, L. Backstrom, D. Huttenlocher, and J. Kleinberg. Mapping the world's photos. In *Proceedings of ACM International Conference on World Wide Web (WWW)* , 2009.

[7] J. Cranshaw, R. Schwartz, J. Hong, and N. Sadeh. The Livehoods Project: Utilizing Social Media to Understand the Dynamics of a City. In *International AAAI Conference on Weblogs and Social Media (ICWSM)*, 2012.

[8] C. Doersch, S. Singh, A. Gupta, J. Sivic, and A. A. Efros. What Makes Paris Look like Paris? *ACM Transactions on Graphics (SIGGRAPH)*, 2012.

[9] P. Glendinning and S. Muthesius. *Tower Block: Modern Public Housing in England, Scotland, Wales, and Northern Ireland*. Yale University Press, 1994.

[10] Ignite. 2012 Social Network Analysis Report. In *Social Media*, 2012.

[11] N. Lathia, D. Quercia, and J. Crowcroft. The Hidden Image of the City : Sensing Community Well-Being from Urban Mobility. In *Proceedings of the International Conference on Pervasive Computing*, 2012.

[12] J. Lindqvist, J. Cranshaw, J. Wiese, J. Hong, and J. Zimmerman. I'm the mayor of my house: examining why people use foursquare - a social-driven location sharing application. In *Proceedings of ACM Conference on Human Factors in Computing Systems (CHI)*, 2011.

[13] K. Lynch. *The Image of the City*. Urban Studies. MIT Press, 1960.

[14] D. Mclennan, H. Barnes, M. Noble, J. Davies, and E. Garratt. *The English Indices of Deprivation 2010*. UK Office for National Statistics, 2011.

[15] S. Milgram. *The individual in a social world: essays and experiments*. Addison-Wesley, 1977.

[16] S. Milgram and D. Jodelet. Psychological maps of Paris. *Environmental Psychology*, 1976.

[17] S. Milgram, S. Kessler, and W. McKenna. A Psychological Map of New York City. *American Scientist*, 1972.

[18] A. Noulas, S. Scellato, R. Lambiotte, M. Pontil, and C. Mascolo. A Tale of Many Cities: Universal Patterns in Human Urban Mobility. *PLoS ONE*, 2012.

[19] D. Quercia, N. Ohare, and H. Cramer. Aesthetic Capital: What Makes London Look Beautiful, Quiet, and Happy? 2013.

[20] C. Smith, D. Quercia, and L. Capra. Finger On The Pulse: Identifying Deprivation Using Transit Flow Analysis. In *Proceedings of ACM International Conference on Computer-Supported Cooperative Work (CSCW)* , 2013.

[21] TfL. London Travel Demand Survey (LTDS). *Transport for London*, 2011.

[22] TheEconomist. More than just a game: Video games are behind the latest fad in management. November 2012.

[23] L. von Ahn and L. Dabbish. Designing games with a purpose. *Communications of ACM*, 2008.

[24] K. Werbach. *For the Win: How Game Thinking Can Revolutionize Your Business*. Wharton Press, 2012.