Playful Cities: Crowdsourcing Urban Happiness with Web Games

DANIELE QUERCIA

It is well known that the layout and configuration of urban space plugs directly into our sense of community wellbeing. The twentieth-century city planner Kevin Lynch showed that a city’s dwellers create their own personal ‘mental maps’ of the city based on features such as the routes they use and the areas they visit. Maps that are easy to remember and navigate bring comfort and ultimately contribute to people’s wellbeing. Unfortunately, traditional social science experiments (including those used to capture mental maps) take time, are costly, and cannot be conducted at city scale. This paper describes how, starting in mid-2012, a team of researchers from a variety of disciplines set about tackling these issues. They were able to translate a few traditional experiments into 1-minute ‘web games with a purpose’. This article describes those games, the main insights they offer, their theoretical implications for urban planning, and their practical implications for improvements in navigation technologies.

Cities in the last decade have been attracting ever more research interest in many disciplines. In particular, the agenda behind the smart cities movement is popular among computer scientists and engineers: new monitoring technologies promise to allocate urban resources (e.g., electricity, clean water, car traffic) more efficiently and, as such, make our cities ‘smarter’. Much of this momentum is being generated by computers and sensors embedded into the built environment, producing data that can be interpreted with respect to existing and new patterns of human behaviour.

In the last few years, I have worked on projects that offer a counterpoint to the dominant efficiency-driven narrative of the smart cities movement. These projects have focused on understanding how people psychologically experience the city. We have considered hypotheses put forward in the 1970s urban sociology literature from commentators, researchers and practitioners such as Jane Jacobs (Jacobs, 1961), Stanley Milgram (Milgram, 1972), and Kevin Lynch (Lynch, 1960), and for the first time, we have been able to test them at scale. We have done so by building two types of crowdsourcing web game: one crowdsources mental images of the city using the example of London (Quercia et al., 2013), and the other crowdsources the discovery of the urban elements that make people happy (Quercia et al., 2014b). Both of these features of the city are important, not least because they have been repeatedly linked to city residents’ wellbeing.

This article focuses on our experience in building those two kinds of web game. The experience offers a new methodology which we believe is able to capture the urban psychological perceptions at the scale of an entire city.

Why Web Games
To begin with, one might wonder why the
urban research communities need web games at all. In the 1960s, scholars started to design experiments that captured the psychological perceptions of city dwellers. They did so by, for example, capturing a city’s psychological maps. Such maps are a subjective representation that dwellers carry around in their heads and are important because of their link to residents’ wellbeing. People generally feel at home in cities whose neighbourhoods are legible and recognizable – that is, cities whose psychological maps require little mental effort in their use and recall, and, as such, are easy to remember and navigate.

In 1972, like Lynch (1960) before him, Stanley Milgram proposed a methodology to capture such psychological maps. He recruited dwellers in New York City, showed them scenes of the city, and tested whether they could recognize where those scenes were: depending on which places were correctly recognized, he drew their collective psychological (recognizability) map (Milgram, 1972).

Not only psychological maps but also visual perceptions have been related to city wellbeing. Consider the idea behind the ‘broken windows’ theory – which suggests that neighbourhood appearances drive the reality of neighbourhood safety (Kelling and Coles, 1998). One broken window leads to another broken window and, in turn, this kind of feedback generates ever more future crime. This theory was not confined to academic circles but made its way into public policy, most notably in New York City. Since visual perceptions were playing a role in public policy, understanding their effects is key. In 1967, Peterson proposed a quantitative analysis of a neighbourhood’s visual appearance (Peterson, 1967). He did so by choosing ten variables that reflected visual appearance (e.g. preferences for the scene, greenery, open space, safety, beauty) and he organized 140 individuals to rate twenty-three pictures of urban scenes in the Chicago metropolitan area for each of those ten variables. He found that beauty and safety are approximately collinear with a preference for a scene, suggesting that the ‘beauty of visual appearance is in fact synonymous with perception of visual pleasure and, hence, desirability of visual appearance’ (Peterson, 1967).

However, the problem with these kinds of experiment is that they take time (hours), are costly (because of paid participants), and cannot be conducted at scale (they involve a few hundreds of participants at most). That is why a number of aspects concerning urban psychological perceptions have been hypothesized, qualitatively studied, but have never been quantitatively tested at scale. To tackle those drawbacks, starting from mid-2012 with various colleagues, our team translated those traditional experiments into 1-minute ‘web games with a purpose’. Such games are said to have ‘a purpose’ because they outsource the not-so-fun activities (e.g. recognizability of urban scenes etc.) to humans in an entertaining way (Von Ahn and Dabbish, 2008). By executing experiments as games, we have been able to harness thousands of human brains. As opposed to study participants, our players happily forked out time for the privilege of being allowed to test their knowledge of the city. To see why they did so, let us focus on the two games, starting with the method of generating a ‘psychological map’.

**UrbanOpticon.org: Crowdsourcing Psychological Maps**

**Psychological Maps**

A geographic map of a city consists of streets and buildings and reflects an objective representation. A psychological map, instead, is a subjective representation that dwellers and participants in the urban scene carry around in their heads. Tourists in a strange city start with few reference points (e.g. hotels, main streets etc.) and then expand the representation in their minds; in short, they slowly begin to build a picture. To see how these
subjective representations matter, consider that, starting from Kevin Lynch’s seminal book *The Image of the City* which was first published in 1960, studies have posited that good imaginability allows city dwellers to feel at home and increase their community wellbeing (Lynch, 1960). People generally feel at home in cities whose neighbourhoods are recognizable. Comfort resulting from little effort, the argument goes, would impact the individual and ultimately collective wellbeing. A case in point is in a large city such as London. Every Londoner has had long attachment to some parts of the city, which bring 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.

Kevin Lynch was one of the first to work formally on these kinds of perception and he created a psychological map of Boston by interviewing Bostonians. Based on hand-drawn maps of what the participants’ ‘versions of Boston’ looked like, he found that a few central areas were known to almost all Bostonians, while vast parts of the city were unknown to its dwellers. More than 10 years later, Stanley Milgram (1972) repeated the same experiment and did so in a variety of other cities (e.g. Paris, New York).

However, the problem with hand-drawn maps is that capturing them takes time, and it is not clear how to aggregate the variety of unique map configurations. One way of fixing this problem is to place a number of constraints on the participants when externalizing their maps. In this vein, Milgram constrained the experiment by reducing it to a simple question: ‘If an individual is placed at random at a point in the city, how likely is he (or she) to know where he is?’ (Milgram, 1972). 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 located. Milgram did setup and successfully run such an experiment in various lecture theatres. Each participant usually spent 90 minutes on the task, and he collected responses from as many as 200 participants for New York City. Since then, such an experimental setup has rarely been replicated. We have recently done so but in the form of a web game (Quercia *et al.*, 2013). As we have heavily borrowed from Milgram’s work, which he titled ‘Psychological Maps’, we have titled ours as ‘Psychological Maps 2.0’.

**Psychological Maps 2.0**

Our crowdsourcing web – available at www.urbanoption.org – 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 less than five months, we collected data from 2,255 users: 739 connecting from London (IP addresses), 973 from the rest of the UK, and 543 outside the UK. A fraction of those participants (287) specified their personal details. The percentage of male/female participants overall was 60/40 and this slightly varied depending on the location: it was 60/40 in London, 65/35 in UK, and 45/55 outside the country. Also, across locations, average age did not differ from the London population, which is 36.4 years old. As for the geographic distribution of respondents, we found a strong correlation between a London region’s population and the number of game participants who lived in the region (with a Pearson correlation coefficient \( r = 0.82 \)). After those users had played the game, we built a collective recognizability map of London based on the fraction of correct answers in each borough.

**Collective Psychological Map of London**

By analysing the recognizability of the five London regions, we found that the general conclusions drawn by Milgram for New York...
hold for London with impressive consistency, adding external validity to our study. Central London was the most recognizable region (about two and a half as many correct placements as for the others) while South London had little cognitive coverage. Figure 1 shows the data as a cartogram of London boroughs. The geometry of the map is distorted based on recognizability scores: Central London dominates, while the South is absent in our respondents’ collective map. We found that there was no relationship between recognizability of a scene and a respondent’s self-reported home location. On the contrary, participants were more likely to recognize scenes in Central London than scenes in their own boroughs. Londoners would answer ‘West London’ and such like when unsure; making the most incorrect guesses for that region – hence a West London response bias.

The recognizability of each region depended on which parts of the city Londoners lived in, but changed depending on whether participants were in UK or not. The number of correct guesses drastically decreased for participants outside London – except for those in Central London. The recognizability in that area remained the same across for all participants: participants outside UK were as good as those within at recognizing scenes in Central London. Hosting some of the most popular tourist attractions in the world, Central London is clearly vividly present in the world’s collective psychological map as anyone travelling around the world soon recognizes.

Taken together, the results suggest two generalizable principles on why people recognized an area. They did so because they were exposed to it (Central London attracts dwell-

Figure 1. A cartogram of London Boroughs. The geographic area is distorted based on Borough’s recognizability.
ers from all over the city), and because the area offered a distinctive architecture (e.g. stadium, tower building) or cultural life (as the central part of East London has been so characterized for generations). We therefore tested the extent to which an area’s recognizability is explained by the area’s exposure to people. In particular, we studied the exposure to users of three social media services and to London underground passengers. We collected 1.2M Twitter messages, 224K Foursquare check-ins, 76.6M underground trips, and 1.3M Flickr pictures in London. Initially with this study, we were aiming to inform social media research in the urban context by establishing which social media data could be used as a proxy for recognizability and exposure, which are key aspects in the study of urban dynamics. It turns out that the answer is complex, suggesting a word of caution for researchers not to take geo-referenced social media data at face value. Not all social media content, even if geo-referenced, can be used in geographic studies. The more the content is associated with physical reality, the more we consider it can be used for geographic studies. For example, tweets generally have little to do with the places in which they are geo-referenced; in contrast, geo-referenced pictures do describe the places in which they are taken. This insight is further supported by our results: visibility is explained more by the exposure to Flickr/Foursquare users than by the exposure to those using Twitter.

In line with Milgram’s experiment with New Yorkers (Milgram, 1972), we found that the acquisition of a mental map is not necessarily a direct process but can also be indirect through, for example, movies. The most distinctive area in London was Blackfriars, and 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’.

Psychological Maps and Socio-Economic Conditions

Starting from Kevin Lynch’s The Image of the City, studies have posited that good imaginability allows city dwellers to feel at home and increase their community wellbeing. With our data we were able to show quantitatively the extent to which the typical Londoners’ collective psychological map tallies with the socio-economic indicators of housing deprivation, living environment conditions, and crime, which we present in figure 2. By correlating the fraction of correct answers in a borough with the borough’s eight socio-economic indicators, we see that

![Figure 2. Correlation between recognizability and deprivation for London Boroughs.](image)
recognizable boroughs enjoy better housing and living environments, and higher incomes. Since previous work related urban wellbeing not only to mental imaginability but also perceptions of beauty, we will now introduce the game that crowdsources a collection of urban visual perceptions.

**UrbanGems.org: Crowdsourcing Urban Beauty**

**Beauty Maps**

Appearances matter in people’s perceptions and experience of their urban environments. Visual features can distinguish one locale from another, and help people perceive a city as unique and recognizable. The ‘broken windows’ theory noted earlier is the most famous theory that links urban appearance to social effects (Kelling and Coles, 1998). Neighbourhood appearances drive the reality of neighbourhood safety: one broken window leads to another broken window and, in turn, to future crime as we implied above.

**Beauty Maps 2.0**

So far the most detailed studies of perceptions of urban environments and their visual appearance have relied on personal interviews and observation of city streets: for example, some researchers relied on annotations of video recordings by experts (Sampson and Raudenbush, 2004), while others have used participant ratings of simulated (rather than existing) street scenes (Lindal and Hartig, 2012). The web has recently been used to survey a large number of individuals. Place Pulse is a website that asks a series of binary perception questions (such as ‘Which place looks safer [between the two]?’) across a large number of geo-tagged images (Salesses et al., 2013).

In a similar vein, we built a web game with collaborators at the University of Cambridge (available at www.urbangems.org). Users are shown ten pairs of urban scenes of London and, for each pair, a user needs to choose which one they consider to be most beautiful, quiet, and happy as we show in figure 3. We

![Figure 3. Screenshot of the crowdsourcing task. The question ‘Which place do you find more beautiful?’ is on top of the two urban scenes. By clicking on that question, the other two on quietness and happiness are made available.](image-url)
chose those three qualities to ground our analysis in classic urban studies, as well as in more popular discussions of ‘city life’.

After 4 months, we had collected data from as many as 3,301 participants: 36 per cent connecting from London (IP addresses), 35 per cent from the rest of the UK, and 29 per cent outside the UK. A fraction of those participants (515) responded to a survey in which they specified their personal details. The percentage of male/female for those participants is 66/34. Their average age was 38 years old (range: 18–77 years old). Compared to the 2001 UK census, our sample was fairly representative, in that, White participants were slightly overrepresented by +6.4 per cent, while participants of Asian descent, Black, Indian, Mixed and Irish were represented in a balanced way. The top country of origin was United Kingdom (65.1 per cent) and the top city of origin was London (40.3 per cent). Professions were quite diverse, the most common being Student, IT Professional, Academic/Scientist, and Architect/Urbanist.

Based on user votes, we found that the strongest affiliation happens to be between beauty and happiness (r = 0.64) as opposed to happiness and quietness (r = 0.29) and beauty and quietness (r = 0.33). Interestingly, the intimate relationship between beauty and happiness was aptly crystallized by Stendhal in his book *On Love*: ‘Beauty is the promise of happiness’ (Stendhal, 1822).

**London: The Aesthetic Capital**

In a way similar to Salesses *et al.* (2013), based on which urban scenes ‘win’ user votes, we assessed the extent to which those scenes are considered quiet, beautiful and happy. We were then able to analyse the scenes together with their ratings using image-processing tools. The goal was to determine what it is about certain neighbourhoods that makes them appear to speak of beauty, quiet and, ultimately, happiness.

In a paper titled ‘Aesthetic Capital: What Makes London Look Beautiful, Quiet, and Happy?’ (Quercia *et al.*, 2014a), we correlated the presence of certain visual features with the quiet, beauty, and happy scores. We saw that the amount of greenery in any given scene was associated with all three attributes and that cars and fortress-like buildings were associated with sadness, where we equated sadness to our measurement for the low end of our ‘spectrum’ of happiness. In contrast, public gardens and Victorian and red brick houses were associated with happiness. Taken all together, our results pointed in the same direction: urban elements that hinder social interactions were undesirable, while elements that increase interactions were the ones that should be integrated by urban planners to retrofit our cities to achieve greater happiness. When we started this work, we were aiming to quantify the aesthetic capital of neighbourhoods. Based on our results, we might hypothesize that aesthetic capital is part of the multi-faceted concept of social capital: urban elements our participants considered beautiful turned out to be those that increase the potential for social interactions.

**Practitioner Reactions**

To gain a perspective from practitioners in the realm of the built environment as a starting point for discussion, we asked three architects (2 males, 1 female, New York City and Barcelona-based) to reflect on our work, and the role of visual elements that our participants associated with beauty, quietness and happiness. They were interviewed in person or over Skype. First, they were given a basic overview of the project and were asked according to their own preference, to either draw or describe a ‘happy scene in London’. Then, they were primed with 10 black-dotted street views indicating visual words (such as those in figure 4) that were highly ranked on happiness. They were then asked to either redraw their scene or to explain how they would change it. They were asked to what extent relationships between visual features...
and people’s reactions played a role in their work, and how visual words could help, or perhaps hinder.

All three indicated that visual words could be useful in providing insight about features affecting people’s reactions. Asked which features stood out, they specifically noted several: the markings related to the importance of green spaces (well-known in their field); the positive effect of pedestrians and cyclists (one architect remarked that the presence of people humanized London and reminded him of Jane Jacobs’s work, noting her emphasis on the presence and activity of people, without referencing a specific scholarly work); suggestions about the variety of colour, types of materials and features (e.g. traditional middle-class housing materials) associated with happiness and beauty; the appearance of many private dwellings, buildings with space for walkways, and buildings whose façades included windows and ornaments; and features that influenced familiarity and comfort (such as common ‘artisanal’ lamp posts) or those which create the ‘London identity’ associated with typical red postboxes, red buses, and red telephone boxes.

They also remarked on potential business or investment opportunities. Knowing what makes people – customers – happy, and then designing for such ‘happiness’ could increase profits, they noted. It could guide investment in urban landscaping by business and property owners. Ratings would indicate appeal to a greater population, and could steer the adding of value to space in a certain way, potentially having consequences for real estate development.

Limitations were also noted. The composition of features, for example, played a role in how urban façades were rated. One noted that visual features would be useful to know, but that she would not rely on them as they might not necessarily reflect the experience of ‘being there’. Additional drawbacks of crowdsourced visual words included those that tended to reflect traditional styles of architecture and, as such, reflected a ‘democratic’ view of English beauty, which may not necessarily match modernist or more forward-looking architecture. Similarly to critiquing ‘democratization of art’ (Komar et al., 1997), it was noted that democratizing architecture may have considerable drawbacks. When asked for a reaction, the third practitioner noted that algorithms might start to pigeonhole certain types of buildings, yet she conceded that algorithms might also be an interesting reference, opening the whole area for discourse and debate.

From our practitioners’ comments, it is clear that cities are not just collections of buildings and views, as one would expect. As pointed out by Hillier, ‘buildings are not just objects, but transformations of space through objects’ (Hillier, 2006); they need to be considered as systems of spatial relations and not as just physical objects. Jacobs, who points out the essential interplay of the ‘bits and pieces’ of a city, warns against bland consensus, and points out that unifying design elements should not be so ubiquitous that they are rendered ineffective (Jacobs, 1961).

Rather than aiming for a ‘consensus’ in what ‘is averagely pleasing’, we aimed to explore the use of crowdsourcing tools and

Figure 4. Visual words for beauty. The black dots are placed next to the visual words that positively correlate with beauty scores.
visual analytics in identifying urban features to which people appear to react. Those initial steps call for new research and a critical debate across a variety of disciplines, including urban informatics, planning and architecture. Ultimately, the vision behind this project is that, with a comprehensive list of aesthetic features at hand, we would be more likely to systematically understand people’s reactions to their surroundings, without striving for a lovable but bland mean.

Discussion

There is strong momentum for making cities smarter, and the ability to identify areas in need could provide real-time information to, for example, local authorities and other agencies with a mandate to make the city better. Those authorities 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. We have seen that one might be able to identify areas in need of intervention also from web games’ data. Web games offer an opportunity to measure the impact of an urban intervention by looking at any difference in participants’ responses before and after the intervention. To do so, one needs to collect data continuously, and that entails effective engagement strategies. The strategies in our games came from research in the area of ‘gamification’ in computer science and included giving points, creating a sense of freshness, and of purpose (Von Ahn and Dabbish, 2008). More specifically, in our two games:

- The purpose of the score was to facilitate the player’s assessment of their performance against previous game rounds or against other players.
- Pictures were chosen randomly to create a sense of freshness and increase replay value. In addition, for experimental sake, randomization reduced biases and led to reliable results, producing a distribution of answers for each picture that was roughly normally distributed.
- To increase the sense of belonging to a community, users could post their scores on Facebook and Twitter after each round. These posts not only fostered a preliminary sense of community but also increased awareness among other social media users.

However, before informing any change in policy, data from games (which are, after all, based on a convenience sample) need to be supplemented by other types of data (e.g. official and curated statistics) and by complementary causal analyses.

Also, the design of web games takes time and, if need be, one could resort to more easily accessible social media data. English neighbourhood deprivation has been related to Twitter topics (Quercia et al., 2012a) and sentiment (Quercia et al., 2012c), and a new way of redefining neighbourhood boundaries has been proposed for Foursquare check-ins (Cranshaw et al., 2012). The effectiveness of such data in the city context comes as no surprise because, after all, a street’s vitality is nowadays captured in the digital layer: street dwellers take pictures and post them on Flickr, and, when they visit places, they share their whereabouts on Foursquare. It is therefore reasonable to assume that there are digital footprints that distinguish lively streets from dead ones (Quercia et al., 2015a), and those footprints – including those coming from mobile phone records (De Nadai et al., 2016) – are ready to be studied in an unprecedented way across temporal and spatial scales.

Conclusion

The main goal of this article has been to illustrate a new methodology that uses computer science tools to replicate social science experiments at scale. In addition to the academic contributions highlighted so
far, our research has resulted in practical implications: for example, it has formed the basis for the creation of new maps, maps where one not only finds the shortest path but also the most enjoyable path (Quercia et al., 2014). With this urban cartography weighted for human emotions, one is not only able to see and connect the shortest segments from point $A$ to point $B$, but also to see the happy path, the beautiful path, the quiet path. In tests, participants found the happy, the beautiful, the quiet paths are far more enjoyable than the shortest one, and this simply involves adding a few minutes to travel time. Participants also recalled how some paths smell and sound. So what if we had a mapping tool that would return the most enjoyable routes based not only on aesthetics but also based on smell (Quercia et al., 2015b) and sound? Being able to capture the multi-sensory layers of what makes for a good city street and more generally a good city is likely to result in innovative digital tools that are aware of people’s perceptions and, more profoundly, they promise a shift in the way we might live our lives in our cities.

REFERENCES


