“...when you’re a Stranger”: Evaluating Safety Perceptions of (un)familiar Urban Places

Martin Traunmueller
ICRI Cities
University College London
Gower Street, London
WC1E 6BT, UK
martin.traunmueller.11@ucl.ac.uk

Paul Marshall
UCLIC
University College London
Gower Street, London
WC1E 6BT, UK
paul.marshall@ucl.ac.uk

Licia Capra
Dept. of Computer Science
University College London
Gower Street, London
WC1E 6BT, UK
l.capra@ucl.ac.uk

ABSTRACT

What makes us feel safe when walking around our cities? Previous research has shown that our perception of safety strongly depends on characteristics of the built environment; separately, research has also shown that safety perceptions depend on the people we encounter on the streets. However, it is not clear how the two relate to one another. In this paper, we propose a quantitative method to investigate this relationship. Using an online crowd–sourcing approach, we collected 5452 safety ratings from over 500 users about images showing various combinations of built environment and people inhabiting it. We applied analysis of covariance (ANCOVA) to the collected data and found that familiarity of the scene is the single most important predictor of our sense of safety. Controlling for familiarity, we identified then what features of the urban environment increase or decrease our safety perception.

CCS Concepts

•**Human-centered computing** → *Empirical studies in collaborative and social computing;*

Keywords

Safety; Perception; Urban; Environment; Crowdsourcing

1. INTRODUCTION

Cities all over the world are being quantified and rated in terms of their livability and life quality for better understanding of citizen satisfaction. One major aspect of urban life quality is the perception of safety in the city: if people avoid feared places, the city’s walkability decreases [11, 27] and this leads to increased motor traffic, which has an impact on the city’s sustainability [27].

Common ways to investigate how people perceive their environment are surveys and semi–structured interviews. In doing so, studies found for instance that visual cues of the built environment, such as broken windows [28], graffiti, abandoned buildings and broken streetlights [2, 19, 22] can trigger a feeling of unsafety. Other cues can be a result of urban planning, or the lacktherof [4, 7].

Other theories still have explored the relationship between perceptions of safety and people, both in terms of *who we are*, and *who we see*. Depending on background and demographic properties, such as age [30], gender [3] and ethnicity [1, 13] people perceive safety differently.

Based on qualitative methods, these works offer semantically rich and detailed insights; however, the cost associated with using these methods makes it difficult to replicate such studies at scale, across different cultures, and over time. To reach a larger number of people, recent research in urban studies and computer science suggests online crowdsourcing as a novel method to gather perceptions, of happiness and safety amongst others, on larger scale. For example, Salesses et al. [20] and Quercia et al. [14] present two randomly selected Google Street View images of a city on a webpage to the user who is asked to select the one appearing more safe or happy based on features of the built environment. Traunmueller et al. [23] use a similar method to evaluate safety perception towards people. In this way, a large amount of data is collected that allows researchers to validate previously defined theories spawn either from architecture or the social sciences.

However, in reality urban places are not made either of the built environment *or* the people who use them, but a combination of the two [24]. In this case, it is not just about how much each of these aspects matter on their own in terms of our safety perception, but even more about how much they matter in interaction with each other. In this paper, we set up an experiment to study this relationship. We built an online crowdsourcing platform similar to Salesses et al. [20] and Traunmueller et al. [23], presenting images of pre–selected types of people overlaid on different urban backgrounds to the user, who is asked to rate them in terms of safety perception and familiarity. We ran the study from October until November 2015, collected and analysed 5425 answers of over 500 participants. On gathered data we used an Analysis of Covariance (ANCOVA) model to draw relationships between participant’s safety perception towards the built environment, the person in the image and participant’s familiarity with the situation. We show that familiarity is the most important single variable that defines if we perceive a place as safe or not. Furthermore we show how various aspects found in the build environment and the people inhabitating it matter to our safety perception while controlling for place familiarity.

The remainder of the paper is structured as follows: next we give an overview of most recent quantitative work on of studies on urban perception. Then we outline the development and deployment of *Streetwise*, an online crowdsourcing platform to gather safety perceptions of urban places, define research question and describe our analysis steps. We then present the results of our study, discuss its limitations, and outline its future directions and opportunities.

2. RELATED WORK

Current work in computer and social sciences aim to establish novel ways to describe urban phenomenas at scale. A method that has recently shown great potential is online crowdsourcing. For instance, *Urbanopticon* [15] presents 360 degree Google Street View images to the user who is asked to guess their geographic location on a map. As 'happy' places are found to be easily recognized by people [8], a collective mental map is drawn with the aim to detect 'happy' places of the city. *Urbangems* [14] suggests an approach to crowdsource perceptions of 'happiness', 'beauty' or 'calmness' of a city based on visual cues of the built environment (e.g., the amount of green space, width of streets) [28]. The online platform shows two random images of Greater London to the user, who is asked to rate images on these attributes. In this way, *Urbangems* aims to identify visual cues of the built environment and the perceived attributes (e.g., happiness, beauty) people attach to them. A follow–up study shows how outcomes can be used to generate 'happy' and 'beautiful' routes through the city [16]. Following the same methodology, *PlacePulse* [20] extends the research into the perception of safety, among other attributes, and not only for one European city, but for different cultures using images of both U.S. American and European cities. Based on these findings, researchers developed *Streetscore* [10], an algorithm that identifies visual cues in Google Street View images using computer vision technology to automate the process. In this way, images for cities all over the world can be scanned, identified by their visual properties and classified as more and less safely perceived environments automatically.

Traunmueller, et al [23] used online crowdsourcing to detect visual cues on people rather than the built environment, that affect others safety perception. The online platform *Streetsmart* shows images of people placed on a white canvas to the user, who is asked to rate images according to his/her safety perception. In this way, *Streetsmart* enabled researchers to evaluate quantitatively established theories based on qualitative work, and to define new ones. Findings suggest for instance, that presence of women, Asian or Caucasian people increase the perception of safety, while men and other ethnicities, such as African–American and Middle–Eastern decrease it. Besides main effects, the study revealed also interactions that showed perception differences within each ethnical group, as for instance the difference in perception of teenagers according to their gender: while teenagers were found to be perceived as safest when being female they were perceived as least safe when being male.

The use of artificial renderings has been discussed in the past in terms of their representativeness [18], but above works show the potential of using online crowdsourcing to gather perception data about safety at scale, whether based on visual cues in an urban environment or on people. However, according to Tuan [24] an urban environment constitutes a *place* shaped by both physical and dynamic factors in interaction with eachother. Studies that used Google Street View images [20, 14] focus on physical elements of the built environment only. In fact, Google Street View images are mostly captured in the early mornings and hence show empty sidewalks, little traffic and generally little activity. They keep out the people dynamics which have great impact on the perception of a place [4, 24]. Studies focussing on people only [23] explore these dynamic factors, but take them out of context. In this paper we develop an online crowdsourcing experiment to explore how built environment, presence of people and familiarity matter to people's safety perception when walking through the city, and what is the relative importance of each when discussed in interaction.

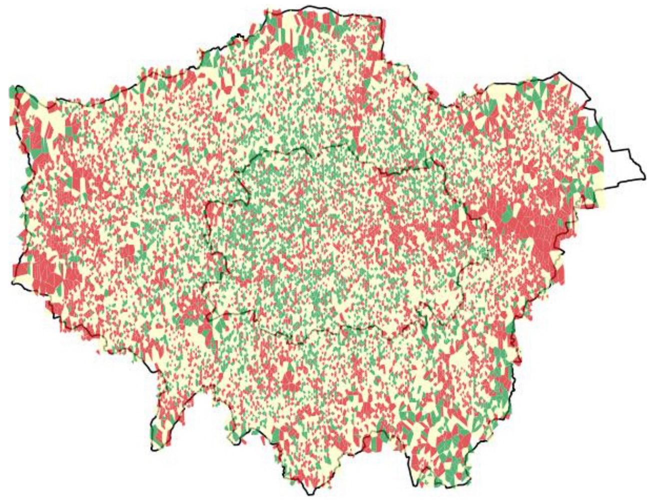

Figure 1: *Streetscore* map for Greater London, showing safety perceptions based on Google Street View images. Green areas are perceived as safe, yellow as neutral and red areas as unsafe.

3. METHOD

We built an online crowdsourcing platform called *Streetwise* to gather safety perception and familiarity data about certain urban situations. Our goal was to understand the impact of place–familiarity, visual properties of built urban environment and the presence of people on our safety perception. In this section, we describe in detail how the platform was built.

3.1 Preparing images

To begin with, we created a pool of images showing various combinations of built environment and people inhabiting it. Images were then used in an online survey to crowdsource the resulting perceived safety by study participants. The selection of background images and the overlayed people were based on previous work. Here we describe the selection process for each of them.

3.1.1 Selecting environments

We ran *Streetscore* [10] for the Greater London area to define a preselection of safe, unsafe or neutral perceived areas based on Google Street View images, as shown in Figure 1. According to their scores we selected 10 images from first, second and third quartile for each type of area. However, as the *Streetscore* algrithm was trained on U.S. American cities it was not sure if it would work well for an European city, such as London. Therefore we ran a small pilot study where we asked people to rate their perceived safety of these 30 images.

We designed an online study to crowdsource their safety perception, where images were rated via Likert–scale in terms of their perceived safety by the participant. We ran the study for two weeks in August 2015, adverted it on social media (Facebook, Twitter) and collected in this way 1590 votes by 53 users who completed a full run of all images. Results of the pilot study confirmed the selection. In the end, we selected the 15 most distinct Google Street View images (5 for each attribute of safe, neutral and unsafe) in terms of mean safety votes and their standard deviations.

Safety perceptions depend on who we are, not just on what we see. Crowdsourcing often fails to reach broad demographics [21]. To ensure representativeness for different demographics, we conducted in addition a qualitative study with selected 15 images following a speak aloud approach. We interviewed 13 people living in

Table 1: People: Selection of images of people used in the study, based on *Streetsmart* findings. Top row shows images of people rated as 'safe', middle row 'neutral' and bottom row 'unsafe' people.

Type	*Images*
safe	
neutral	
unsafe	

London, between the age of 21–52 years, including 8 female and 5 male. Their ethnicities included Caucasian, African–American, Indian, Asian and Arab, covering all main ethnical groups in London. Each session lasted for 30 minutes, answers were audio recorded and transcribed for analysis. After transcribing audio recordings and safety votes, our data showed similar results to the findings of the quantitative part.

Overall, safe streets were found in middle and upper–class neighbourhoods of West London, such as Chelsea and Notting Hill, and unsafe streets were found in working–class neighbourhoods of East London, such as Hackney and Tower Hamlets. Neutral neighbourhoods were found to be in middle–class areas of Islington and Lambeth.

3.1.2 Selecting people

Having defined a selection of urban backgrounds, we next selected images of people to be overlaid on the top of them. We used the findings of [23] to define sets of safe, unsafe or neutral perceived types of people. Therefore we selected 5 images from first, second and third quartile, according to the safety scores. All images used in the pilot study were selected from online repositories under a creative commons license.

Overall, people that are perceived as unsafe included mostly middle – aged men, while safe perceived people mostly included female and elder people. Neutral people included mostly young and elder men (see Table 1).

3.1.3 Overlaying people on backgrounds

Having selected background images and images of people, we were now able to create various combinations overlaying each of them with one another, as shown in Figure 2. Images of people were scaled to appear at a distance of 6–7 meters to the participant and placed in the matching perspective angle on each background image. The distance was chosen as in real situations it enables us to see enough visual detail in other people to assess whether to be fearful or not [5]. We used every image of a person on every image of background, resulting in a total set of 225 images to be used in our study.

Figure 2: Overlaying people on environments: Image examples for a safe (top row) and unsafe (bottom row) environment with safe (left column) and unsafe (right column) perceived people overlayed, as used in the study.

3.2 Streetwise

With our images at hand, we next built the *Streetwise* website to present them to an audience and gather safety perception data. The entry page to the website asked crowdworkers to provide basic background information on who they are, such as their *gender, age* and *ethnicity* and where they come from. Then the safety perception survey followed (see Figure 3). We asked crowdworkers to complete a run of 30 safety evaluations using a slider built as a 100 point Likert–scale (1 = very unsafe, 100 = very safe). With a similar slider, they were asked to indicate their familiarity based on the visual properties of the situation depicted in the image (1 = very unfamiliar, 100 = very familiar). In addition, we asked the user to comment on the voting decision in a commenting box.

User comment, in connection with vote and time–stamp for each click, gave us a good indication about seriousness of user feedback and hence enabled us to detect meaningless data we could exclude from the study. In addition, we included images in the study run showing situations that were expected to raise the fear level of the user and hence would be expected to be rated with low perceived safety scores. Such images showed, for instance, scenes of people carrying guns or wearing scary masks. If participants rated them with high safety scores they were considered as untrustworthy and excluded from the study. Including these images to our image pool, we used 240 pictures that were randomized, so to avoid the same crowdworker seeing the same image twice, and to obtain roughly the same number of answers for each image in each category.

3.3 Data collection

We obtained ethical approval to conduct the study in September 2015. We then deployed *Streetwise* from October 2015 until end of November 2015 and advertised on various social media platforms (Facebook – advertisment, page and group postings –, LinkedIn, Twitter and Reddit). Voluntary users could share the webpage link via embedded Facebook–Like and Tweet buttons. In parallel, we used online marketplace Amazon Mechanical Turk (AMT) to recruite crowdworkers. As our images show an urban environment of a western metropolis, we focussed on turkers from Europe and

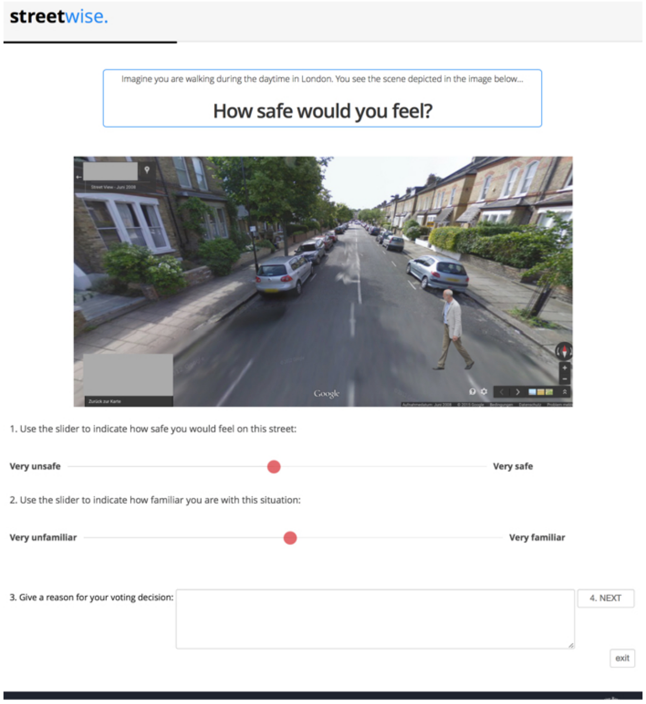

Figure 3: Web–based user–interface of *Streetwise*, showing one image of a person in an urban environment at a time. The user is asked to rate his/her perception of safety and familiarity with the situation according to the image on the sliders, and to give reasons. By pressing the 'Next' button a new image is presented; by pressing 'Exit' the study can be ended at any time.

the U.S. who are expected to be more familiar with these environments.

Throughout these two months, we were able to collect 1290 votes from 202 users recruited from social media, and 4876 votes from 400 AMT crowdworkers. After removing suspicious data (e.g., because of votes given too quickly) from both, we ended up with 5452 votes from 502 participants in total (1130 votes from 173 users recruited from social media, and 4479 votes from 337 AMT crowdworkers). The ratio between male (56%) and female (44%) participants was almost even, with most participants aged between 21–40 years (64%), followed by 41+ year olds (24%) and 0–20 year olds (12%). For ethnicity we found a high majority of Caucasian (78%), followed by African–American (8%) and Asian (6%) people.We merged the two datasets as received from AMT crowdworkers and our participants that have been recruited over social media and normalized safety and familiarity scores to a range between 0 and 1.

3.4 Research Question

Using the collected data we aim to answer the following question:

- How do built environment, people presence and familiarity affect people's safety perception when walking through the city and what is the relative importance of each when combined?

3.5 Analysis

Since our data includes both categorical (type of built environment, type of person in the image) and continous (familiarity score) variables, we used an Analysis of Covariance (ANCOVA) with planned contrasts and post hoc analysis, using the Tukey correction [25] for multiple comparisons to detect significant effects. The planned contrasts take the most common demographic type of people found on London streets (a middle–aged Caucasian female, facing towards the user) as baseline to compare against other demographics. Looking at the frequency distribution of ratings we observed that the data is highly skewed showing a long tail distribution: most people feel very safe in relation to the presented urban situations, whereas only some feel unsafe. This is a common situation in multi–factorial experiments in human–computer interaction (HCI) that work with Likert responses. To transform our data so that we could build an ANCOVA model, we used the *Aligned Rank Transform* [29] for non–parametric factorial data analysis, which aligns the data in a pre–processing step before applying averaged ranks. With transformed data, we were then able to conduct our analysis.

Table 2: Table showing various models of various combinations of variable groups and their $\mathrm{adj.R2}$, with p–value: 0.001 * 0.01 ** 0.05 *.**

Model	*contained Variable Groups*	$adj.R2$	
Model A	Familiarity	0.38	***
Model B	Built Environment	0.10	***
Model C	People Variables	0.08	***
Model D	Built Environment + People Variables	0.17	***
Model E	Familiarity + Built Environment	0.42	***
Model F	Familiarity + People Variables	0.41	***
Model G	Familiarity + Built Environment + People Variables	0.45	***

4. RESULTS

4.1 Overall safety perception of places

We started our analysis by building a set of models consisting of our main variable groups of *Built Environment*, *People Variables* and *Familiarity* first seperately and then combined with each other to detect their influence on overall safety perceptions. For each of them, we took the overall safety score as dependend variable and variables of built environment safety level, people (such as age, gender, ethnicity and facing direction) and familiarity as independent variables. In Table 2 we show each model's performance.

Looking at the results it becomes clear that *Familiarity* with a place is the single most important predictor (Model A, $adj.R2 = 0.38$), with *Built Environment* (Model B, $adj.R2 = 0.10$) and *People Variables* (Model C, $adj.R2 = 0.08$) having less impact. These results suggest that people mainly evaluate if an urban place is safe or not by their own background and experience, defining familiarity. The actual visual properties of the built environment or even the people variables almost do not matter. However, by combining the three groups we were able to improve these results so that the model containing all variables showed best performance (Model G, $adj.R2 = 0.45$).

In Table 3 we present the estimates for each variable of Model G. Results suggest that *Familiarity* ($F(1, 1564) = 2839.82, p < 0.001$) is the single most important factor for safety perception. The more familiar people are with a place, the safer they feel ($\beta = 0.58$), supporting findings from [12]. Other properties found in *Built Environment* and *People Variables* showed less effects:

Table 3: Estimates for each of our discussed variables and containted factors of Model G, compared to defined baseline (middle–aged Caucasian female, facing towards the user on a neutral background), with p–value: 0.001 * 0.01 ** 0.05 *.**

	Variable	*contained Factors*	β	F–Value	
Familiarity	Fam		0.58	2839.82	***
Built Environment	Bi	safe	0.16	147.183	***
		unsafe	–0.32		***
Gender	Gi	male	–0.18	62.046	***
Age	Ai	elder	0.21	41.132	***
		teenager	0.06		*
Ethnicity	Ei	African–American	–0.32	14.158	***
		Asian	–0.28		***
		Middle–Eastern	–0.12		***
Facing	Fi	away from me	0.32	39.523	***
		not aware of me	0.08		**

- *Built Environment* Bi – Built environment ($F(2, 162) = 147.183, p < 0.001$) was found to affect people's safety perception more when being considered as unsafe ($\beta = -0.32$) than safe ($\beta = 0.16$), suggesting that people are affected by their built surroundings more negatively than positively.
- *Gender* Gi – The gender of the person in the image was found to have significant impact ($F(1, 34) = 62.046, p < 0.001$), decreasing safety perception when men were shown in the image ($\beta = -0.18$) compared to women.
- *Age* Ai – The age of the person in the image was found to have significant impact ($F(2, 45) = 41.132, p < 0.001$), increasing people's safety perception when elder people were shown in the image ($\beta = 0.21$) compared to middle–aged people. Presence of teenagers showed less positive impact ($\beta = 0.06$).
- *Ethnicity* Ei – For ethnicity ($F(3, 23) = 14.158, p < 0.001$) we found that compared to Caucasian, every other ethnicity decreased safety perception significantly, such as African–American ($\beta = -0.32$), Asian ($\beta = -0.28$) and Middle–Eastern ($\beta = -0.12$). These findings are in fact surprising, as Traunmueller et al. [23] found especially the presence of Asian people increasing the perception of safety, when being discussed without environmental context.
- *Facing Direction* Fi – Facing direction of people showed significant effect ($F(2, 44) = 39.523, p < 0.001$) on safety peception of people. We found especially images with people facing away from the user ($\beta = 0.32$) to increase the user's safety perception, compared to when people looking at the user. Images with people seen from the side ($\beta = 0.08$), such as by–passers, showed less effect.

Results of our overall model suggest that familiarity with a place is the single most important variable that affects our safety perception – the more familiar we are with a place, the safer we feel. Visual properties of the environment has less impact. However, familiarity is very subjective, based on personal background and experience. We do not know how it is being interpreted by people. Therefore the question arises what affects people's safety perception when being in a more or less familiar place?

4.2 Safety Perception in (un)familiar Places

To see how *Built Environment* and *People Variables* impact on our safety perception, once we control *Familiarity*, we devided our images in terms of familiarity scores. Using the standard deviation of familiarity scores ($1st.Qu. = 0.43$, $3rd.Qu. = 0.85$) as breaks, we defined three groups of familiar ($n = 1545$), neutral ($n = 2368$) and unfamiliar ($n = 1539$) rated images and built three ANCOVA models relating to place–familiarity. In Table 4 we show for each model $adj.R2$ value and the β–estimate for those factors we found significant.

Table 4: Estimates for each of our discussed variables and containted factors, compared to defined baseline (middle–aged Caucasian female, facing towards the user on a neutral background), with p–value: 0.001 * 0.01 ** 0.05 *.**

Image Group	$adj.R2$	*Variable*	*contained Factors*	β	
familiar	0.07	Bi	unsafe	–0.33	***
		Gi	male	–0.20	***
		Ai	elder	0.27	***
			teenager	0.10	*
		Ei	African–American	–0.60	***
			Asian	–0.66	***
		Fi	away from me	0.63	***
neutral	0.12	Bi	safe	0.23	***
			unsafe	–0.41	***
		Gi	male	–0.30	***
		Ai	elder	0.29	***
		Ei	African–American	–0.25	*
		Fi	away from me	0.24	*
			not aware of me	0.14	*
unfamiliar	0.28	Bi	safe	0.33	***
			unsafe	–0.40	***
		Gi	male	–0.19	**
		Ai	elder	0.32	**
		Ei	African–American	–0.43	***
			Middle–Eastern	–0.26	**
		Fi	away from me	0.38	**
			not aware of me	0.16	*

While $adj.R2$ values are generally rather low, they clearly show that the less familiar we are with a place, the more it matters how we perceive built environment and people around us to our safety perception, with an $adj.R2 = 0.28$ for unfamiliar, and 0.07 for familiar places. In more detail we found for:

- *Built Environment* Bi – We found that, when being in a familiar situation, only unsafe perceived built environments have an effect, decreasing people's safety perception ($\beta = -0.33$). Being in an unfamiliar situation, both safe and unsafe perceived built environments are affecting us: positively, when being in a safe environment ($\beta = 0.33$) and negatively, when being in an unsafe environment ($\beta = -0.40$).
- *Gender* Gi – We found that the presence of men decreases people's safety perception, compared to the presence of women, no matter how familiar they are with a situation. While having a similar impact in familiar ($\beta = -0.20$) and unfamiliar ($\beta = -0.18$) situations, men decrease people's safety perception most in neutral environments ($\beta = -0.30$).
- *Age* Gi – We found that, compared to the presence of middle–aged people, elder people increase people's safety perception no matter how familiar they are with the situation. We found the positive impact to be bigger in unfamiliar ($\beta = 0.32$) than in familiar situations ($\beta = 0.27$). In addition, the presence of teenagers showed a minor positive impact in familiar situations ($\beta = 0.10$).
- *Ethnicities* Ei – Other ethnicities than Caucasian were found to decrease safety perceptions in all three familiarity–situations. While African–American people decreased safety perception in all three situations (familiar ($\beta = -0.60$), neutral ($\beta = -0.25$), unfamiliar ($\beta = -0.43$)), the presensce of Asian people significantly decreased people's safety perception only

in familiar situations ($\beta = -0.66$), and Middle–Eastern people only in unfamiliar situations ($\beta = -0.26$).

- *Facing Direction Fi* – People facing away increase people's safety perception and have more effect in familiar ($\beta = 0.66$) than in unfamiliar ($\beta = 0.38$) or neutral situations ($\beta = 0.24$).

In summary, we see how properties found in the urban environment affect our safety perception differently according to our familiarity with an urban place. While built environment is affecting our safety perception both positively and negatively in unfamiliar situations, we found that it only affects us negatively in familiar situations. Furthermore, we found that some demographic properties of people we encounter on the streets affect us in familiar and unfamiliar situations, others do not. For instance, while gender of people matters most in neutral situations, people's age matters most in unfamiliar situations. People's ethnical background was found to impact depending on the grade of familiarity with a situation differently: while African–American people affected safety perceptions in all three situations, presence of Asian people mattered only in familiar and Middle–Eastern people only in unfamiliar situations.

These results show the complexity of human perception in an urban environment based on visual properties and people's background. Next we will discuss these findings, their limitations and future directions of this work.

5. DISCUSSION AND FUTURE WORK

In this paper, we explored quantitatively the relationship between several visual properties found in the urban environment, place–familiarity and the resulting perception of safety. To do so, we have taken findings from prior work on the topic, discussing these visual properties as seperate entities, and built an online platform to crowdsource the perception of safety when being combined with each other. We found that familiarity of the place is the single most important predictor of our perception of safety. We then controlled for familiarity, and identified what other features (e.g., presence of certain people, built environments) increase or decrease our safety perception. Our findings show that in unfamiliar situations, for example, built environment can affect us positively, and that the presence of elder people increases our sense of safety, while the presence of men decreases it.

5.1 Limitations

However, using online images in research brings up a number of limitations in terms of representativeness [18], especially when they are used to crowdsource safety perceptions of the urban environment.

Built Environment – The question arises whether three dimensional urban space can be represented on two dimensional images on a computer screen. Besides its visual properties, urban space is being defined by many other variables that we perceive subconciously through other senses, such as hearing and smelling [17]. That is, we experience a city not just through single images one by one, but through movement, developing a sense of place over time [6]. Our platform does not capture any of these, but focuses on the visual characteristics of the urban environment only, and therefore has to be seen as a first step that leaves space for future research to add other properties to it. Furthermore, cities differ all over the world depending on history and culture of their population. Urban design principles of an organically grown European city are significantly different from an U.S. American city for instance, resulting in differences in urban scale and architectural properties of building facades. In our study, we focussed on London; we do not know whether the outcome would differ when running a smilar study for another city and how. However, it would not be costly to repeat the study focussing on other cities as the method allows it.

People – Based on previous work [23], we discussed in this paper a number of variables defining visual properties of people, such as age, gender, ethnicity and people's facing direction. The way people look is very personal and can differ in a variety of ways. There are many small visual details that might have a big effect on safety perceptions that have not been covered by past work and could be included in future work. Furthermore, the number of people matters on how they are perceived by others [9]. Cities are densely populated areas. When walking through a city, we encounter mostly not just one person at a time, but several people. In this study, we discussed the influence of one person at a time only, and we leave it to future work to investigate the impact of various compositions of people on safety perception.

Crowdsourcing – Safety perceptions are very personal, defined not only by *what we see*, but also by *who we are*. Previous work suggests that crowdsourcing methods might lead to results that are biased by their crowd [23], especially when using Amazon Mechanical Turk [21]. Demographic data gathered from our study participants reflects these circumstances, as we received feedback mostly from Caucasian, middle–aged U.S. Americans. Therefore our results reflect the opinion of this crowdworker demographics only.

5.2 Future Work and Opportunities

For next steps of reserach, we will first focus on the *Built Environment*, including urban environments other than London. With the aim to detect possible differences in safety perception due to varying urban design principles, we will select several cities that are represented on Google Street View, such as Beijing, New York City and Sao Paolo, and include them in the study. As next step, we will then include different compositions of *People*, such as groups and crowds, instead of showing a single person at a time, to detect safety perceptions related to the number of people. As in the study for London, we will include a representative demographic breakdown for people inhabiting each city. In parallel to the expansion of our study, we will also work on different approaches of *Crowdsourcing* other than using AMT, to engage and motivate specific crowds. One possibility is to explore gamification approaches [26], but also social cause and intrinsic rewards with different reward strategies possibly attracting different demographics.

Apart from this, our work offers opportunities to understand urban environments and how they are perceived by their population. So called 'soft data' about how people perceive the urban environment, especially if they feel safe or not, has been difficult to grasp on a large scale. At the same time, these perceptions have great impact on sustainability of a city and urban life quality of its population: if people avoid feared places, the city's walkability decreases [11, 27] leading to less social interaction among the population and more motor traffic within the city [27]. The method proposed in this paper can be used to harness this 'soft data' more easily, and it is thus a powerful instrument in the hands of social and urban scientists to develop and evaluate complex urban theories at large scale and at little cost. The findings emerging from the use of such method can then be used in practise to build tools on top of them, to the benefit of different stakeholders: administrators can use them to intervene in community development; city planners can use them to guide design principles; and developers can use them to build applications to support urban walking for instance, fostering the sense of communities and hence contributing to urban life quality.

6. ACKNOWLEDGEMENTS

The research leading to these results is part of a PhD that is being funded by the Intel Collaborative Research Institute on Sustainable Connected Cities (ICRI Cities). The use of the Streetscore–Algorithm for Greater London in this study happened in collaboration with the MIT Medialab.

7. REFERENCES

[1] K. Day. 1999. Embassies and sanctuaries: women's experiences of race and fear in public space. *Environment and Planning,* D, 17 (1999), 307–328.

[2] H.. Doeksen. 1997. Reducing crime and the fear of crime by reclaiming New Zealand's suburban street. *Landscape and Urban Planning* 39, 2-3 (1997), 243–252.

[3] M. Felson and R.V. Clarke. 1998. *Opportunity Makes the Thief: Practical theory of crime prevention*. Home Office.

[4] J. Gehl. 2010. *Cities for People*. Island Press.

[5] E.T. Hall. 1966. *The hidden Dimension*. Anchor Books.

[6] H. Heft and J. L. Nasar. 2000. Evaluating environmental scenes using dynamic versus static displays. *Environment and Behavior* 32, 3 (2000), 301–322.

[7] J. Jacobs. 1961. *The Death and Life of Great American Cities*. Random House Inc.

[8] K. Lynch. 1960. *The Image of the City*. MIT Press.

[9] S. Matei, S.J. Ball-Rokeach, and J. Linchuan Qiu. 2001. Fear and Misperception of Los Angeles Urban Space – A Spatial-Statistical Study of Communication-Shaped Mental Maps. *Communication Research* 28, 4 (2001), 429–463.

[10] N. Naik, J. Philipoom, R. Raskar, and C.A. Hidalgo. 2014. StreetScore - Predicting the Perceived Safety of One Million Streetscapes. In *CVPR Workshop on Web-scale Vision and Social Media*.

[11] J.L. Nasar and B. Fisher. 1993. Proximate physical cues to fear of crime. *Landscape and Urban Planning* 26 (1993), 161–178.

[12] J. L. Nasar. 1994. Urban design aesthetics: The evaluative qualities of building exteriors. *Environment and Behavior* 26, 3 (1994), 377.

[13] R. Pain. 2001. Gender, Race, Age and Fear in the City. *Urban Studies* 38, 5–6 (2001), 899–913.

[14] D. Quercia, N. O'Hare, and H. Cramer. 2014a. Aesthetic Capital: What Makes London Look Beautiful, Quiet, and Happy. In *CSCW'14*.

[15] D. Quercia, J.P Pesce, V. Almeida, and J. Crowcroft. 2014b. Psychological Maps 2.0: A web gamification enterprise starting in London. In *Proc of WWW'13*.

[16] D. Quercia, R. Schifanella, and L.M. Aiello. 2014. The Shortest Path to Happiness: Recommending Beautiful, Quiet, and Happy Routes in the City. In *Proc of HT'14*.

[17] D. Quercia, R. Schifanella, L. M. Aiello, and K. McLean. 2015. Smelly Maps: The Digital Life of Urban Smellscapes. In *Proc of ICWSM'15*.

[18] G. Rose. 2007. *Visual Methodologies: An Introduction to the Interpretation of Visual Materials*. SAGE Publications.

[19] C. E. Ross and J. Mirowsky. 1999. Disorder and decay: the concept and measurement of perceived neighborhood disorder. *Urban Affairs Review* 34, 3 (1999), 412–433.

[20] P. Salesses, K. Schechtner, and C.A. Hidalgo. 2013. Image of The City: Mapping the Inequality of Urban Perception. *PLoS ONE 8 (7): e68400.Doi:10.1371/journal.pone.0068400* 8, 7 (2013).

[21] S. Sen, M.E. Macademia Team: Giesel, R. Gold, B. Hillmann, M. Lesicko, S. Naden, J. Russell, Z.K. Wang, and B. Hecht. 2015. Turkers, Scholars, "Arafat" and "Peace": Cultural Communities and Algorithmic Gold Standards. In *Proc of CSCW'15*.

[22] W.G. Skogan. 1999. *Measuring what matters: crime, disorder and fear. Measuring what matters: proceedings from the police research institute meetings*. R. H. Langworthy (Ed.). National Institute of Justice, Washington, DC.

[23] M. Traunmueller, P. Marshall, and L. Capra. 2015. Crowdsourcing Safety Perceptions of People: Opportunities and Limiations. In *Proc of SocInfo 2015*.

[24] Y.F. Tuan. 2001. Space and Place: The Perspective of Experience. (2001).

[25] J. W. Tukey. 1949. Comparing Individual Means in the Analysis of Variance. *Biometrics* 5, 2 (1949), 99–114.

[26] L. von Ahl and L. Dabbish. 2008. Designing games with a purpose. *Communications of the ACM 51* 51, August 2008 (2008), 58–67.

[27] M. Warr and C.G. Ellison. 2000. Rethinking social reactions to crime: personal and altruistic fear in family households. *Amer. J. Sociology* 106, 3 (2000), 551–578.

[28] Q. Wilson, J and G. L. Kelling. 1982. "Broken Windows: The police and neighborhood safety". *Atlantic Monthly* 249, 3 (1982), 29–38.

[29] J.O. Wobbrock, L. Findlater, D. Gergle, and J.J. Higgins. 2011. The Aligned Rank Transform for Nonparametric Factorial Analyses Using Only ANOVA Procedures. In *Proc of CHI'11*.

[30] R. Zako. 2009. Young People's Gatherings. *Urban Public Realm* 66 (2009).