A Quantitative Approach to Evaluate and Develop Theories on (Fear of) Crime in Urban Environments

Martin Wolfgang Traunmueller

May 4, 2017

I, Martin Wolfgang Traunmueller, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the work.

Abstract

Well established work in criminological, architectural and urban studies suggests that there is a strong correlation between *crime*, perceived safety, the *fear of crime*, and the presence of different demographics, the *people dynamics*, in an urban environment. These studies have been conducted primarily using qualitative evaluation methods, and are typically limited in terms of the geographical area they cover, the number of respondents they reach out to, and the temporal frequency with which they can be repeated. As cities are rapidly growing and evolving complex entities, complementary approaches that afford social and urban scientists the ability to evaluate urban *crime* and *fear of crime* theories at scale are required.

In this thesis, I propose a combination of methodologies following a data mining and crowdsourcing approach to quantitatively validate these theories at scale, and to support the exploration of new ones. To relate *people dynamics* to *crime* quantitatively, I first analyse footfall counts as recorded by telecommunication data, and extract metrics that act as proxies of urban crime theories. Using correlation and regression analysis between such proxies and crime activity derived from open crime data records, the method can help to understand to what extent different theories of urban crime hold, and where.

To relate *people dynamics* to *fear of crime* quantitatively, I then built two image–based online crowdsourcing platforms to investigate to what extent online crowdsourcing can be used to gather safety perceptions about urban *places*, defined by the combination of built environment and the people inhabiting it. As existing theories suggest that knowing who the respondents are is crucial for understanding safety perceptions, I also gathered their demographic background information to discuss their perceptions accordingly. I applied analysis of variance (ANOVA) and covariance (ANCOVA) to these data. The method can help to understand what visual properties based on people demographics relate to safety perception in the built environment.

Acknowledgements

Pursuing a PhD is a serious commitment over a long time period, full of ups and downs as are most things in life. Such an undertaking can not be achieved alone, but requires people giving feedback, advise, inspiration, food for thought or simply just pleasure. There are a number of people I want to thank for their support throughout these years, without whom I do believe this work would not have been possible in that way.

First of all, I want to thank my supervisor Licia Capra, who had the pleasure of introducing an architect to the computer science community – not the easiest task, I admit. With her endless patience, sharp feedback and necessary guidance, Licia gave me the opportunity to expand my knowledge and to grow as researcher between the two communities over the years. In doing so, she was supported by Paul Marshall, my second supervisor, whom I want to thank for his constructive feedback, helpful discussions and advise; this work would not have been possible without them. I also want to thank Ava Fatah, my former MSc supervisor at The Bartlett, academic mentor and friend, for her advise and inspiring discussions, that started even before the PhD and somehow guided me into this exciting field of research. I'm also greatful to Giovanni Quattrone, former Post Doc researcher at the Department of Computer Science and a friend of mine, who especially in the first year of my PhD, was very helpful with his constructive feedback. Furthermore, I want to thank the ICRI as very generous sponsor, enabling me as researcher focussing purely on the work without any distractions. At the same time, it offered me a dynamic lab atmosphere with fellow PhD's and Post Docs, making my life as PhD student less lonely. In particular, I want to thank at this point Sarah Gallacher and Martin Dittus for their support, advise and helpful discussions.

I'm very grateful to my family, especially to my girlfriend Isin and my mom and dad, for their unconditional support, their great interest and understanding throughout the years. By now, they know a lot more about the relationship between *people dynamics* and (fear of) crime in cities.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Overview

In modern society we are experiencing two phenomena: there is a rapid population shift of people moving from rural areas into urban environments, with an annual growth of 60 million new city dwellers every year (WHO, 2010). In parallel, we are experiencing, as in the UK for example, a steady rise in police recorded crime activities over the last years (Flatley, 2015), focusing especially on densely populated areas (Jansson, 2006; Bettencourt et al., 2007; Glaeser, 1999). To face this situation, cities have made an effort over the last years to lower crime, resulting in an increased interest in crime research. The main goal of this research has been to establish the relationships between crime and population (Tan and Haining, 2009; Song and Daqian, 2013), and between crime and the built environment (Hillier and Sahbaz, 2009; Sahbaz and Hiller, 2007; Wolfe and Mennis, 2012).

Besides crime, fear of crime has become an increasing problem for the broad population (Brown and Polk, 1996; Oc and Tiesdell, 1997) that has to be included into the discussion. In contrast to actual crime activity, fear of crime describes a perception, such as a lack of feeling of safety by the urban population in the environment, that stands in close relation to crime activity (Doran and Burgess, 2012). However, supported by newsbroadcasting through modern media, such as television and online social networks (Matei et al., 2001), fear of crime is not restricted to space and time, and hence is not limited to appear in areas of high crime activity only. These circumstances lead to serious problems for a city's inhabitants and its government as fear of crime has great impact on the quality of urban life (Pacione, 2003): for example, as people avoid feared places, the city's walkability gets limited (Nasar and Fisher, 1993; Warr and Ellison, 2000) and hence leads to increased motor traffic, impacting the city's sustainability (Warr and Ellison, 2000).

Research that tries to understand the complex dynamics behind crime and fear of crime in urban environments suggests that there is a relationship between properties of the city's population (*people dynamics*), such as people's age, gender and ethnicity, and (fear of) crime. Most work is grounded in the same theories, primarily Jacobs (1961) and Newman (1972), who suggest opposite approaches to the design of urban space to gain control over who is inhabiting it, and hence to take advantage of the resulting social control of (fear of) crime: by including or excluding different demographic population groups through urban design, they describe *people dynamics* that attract crime activity to, or repel it from, an area and impact fear of crime perception at different times.

Following their work, the discussion can be split up in two different schools of thought. Jacobs (1961) defines urban population as 'eyes on the street', a natural policy mechanism that supports urban safety through 'natural surveillance'. She claims that 'eyes on the street', belonging to concerned urban population, or "natural proprietors" of the street, are necessary for a street or an urban place to be safe. Her work hypothesises that an open and mixed–use environment supports this concept by enabling diversity and activity within the population using the area at different times, leading to more safety.

While Jacobs suggests that a high diversity among the population and a high ratio of visitors are contributing to an area's safety, Newman (1972) argues the opposite. He claims that diversity and a high mix of people create the anonymity needed for crime to take place. According to his theory, Newman suggests that a clear definition of public, semi–public and private space in a low–dense and single–use urban environment creates a 'defensible space' that is needed to support safety. Newman further hypothesises that low population diversity, low visitor ratio and a high ratio of residents contribute to an area's safety.

1.2 Problem Statement

Each theory has been evaluated, and indeed supported, by means of qualitative research methods that enable in–depth and semantically rich investigations into the reasons behind certain phenomena relating to *people dynamics* and (fear of) crime. However, these methods are very expensive and time–consuming to run so that studies are usually restricted to a rather small number of people (relative to the overall urban population) and constrained to geographic areas (e.g., a neighbourhood): In fact, Jacobs' findings are based on observations of individual people in American cities, mostly from New York City neighbourhoods, focussing on her home area of Greenwich Village only. Furthermore, they are almost never repeated over time to observe potential changes. It is thus difficult to understand which of these theories hold – where and when.

Over the years, these theories allowed the development of urban design frameworks, such as *Crime Prevention Through Environmental Design* (CPTED) (ICA, 2000). As such, they have significant impact on the design and use of modern cities all over the world. Therefore it is important to develop an understanding of the relationship between *people dynamics* and (fear of) crime. To complement traditional studies, I propose in this thesis a novel approach, using ICT (information and communication technology) based methodologies in this domain to study the subject at temporal and spatial scale.

1.3 Hypotheses

The rise of modern technology and an increasing availability of open data offers new opportunities for researchers to develop methodologies to study crime and fear of crime on large scale. Besides offering a novel perspective on the topic, outcomes of these studies can be used in triangulation with traditional methods to support their findings on one hand, and to allow researchers to define new questions on the other.

As the definition of the research question is dependent on the availability of data, there are two underlying reserach hypotheses in this thesis:

1st Research Hypothesis: *We can extract features from ready available CDR (Call Detail Records) data to validate urban crime theories at scale.*

2nd Research Hypothesis: *We can use online crowdsourcing to collect perception data and use it to validate and develop fear of crime theories at scale.*

1.4 Contributions

With this research, I aim to introduce methodologies that support urban crime and fear of crime research in the communities of computer science, urban and social / criminological studies. My contributions include:

- A method to support researchers in validating urban crime theories at scale, based on the analysis of mobile phone data.
- A method to support researchers in validating and expanding fear of crime theories towards people and places at scale, based on the analysis of perceptual data collected via online crowdsourcing.

Offering opportunities to study urban crime and fear of crime "at scale", these methods enable researchers to cover larger geographic areas at fine level of spatio–temporal granularity than what qualitative methods afford, as for instance on city–level. As qualitative studies are time and resources–consuming due to the collection and analysis of data by researchers, they impact the area and frequency they are able to cover. A quantitative approach, as suggested in this thesis, reduces the time factor dramatically, as data, such as mobile phone data, is already there, crowdsourced data is collected within hours / days for a large geographic area, such as a city, and the analysis is done within seconds through algorithms.

In this thesis, I applied these methodologies to a specific city (London) as case study, to investigate well established theories (Jacobs, 1961; Newman, 1972) on the relationship between *people dynamics* and (fear of) crime in the urban environment.

1.5 Publications

Publications resulting from this thesis:

	Contribution	Publication
Chapter 4	A methodology that uses CDR data to evaluate urban crime theories at scale.	Traunmueller, M., Quattrone, G. and Capra, L., Mining Mobile Phone Data to Investigate Urban Crime Theories at Scale, *In Proceedings of the 6th International Conference on Social Informatics (SocInfo 2014)*, November 2014. Traunmueller, M., Quattrone, G., Capra, L. and Fatah gen. Schieck, A., A Data Mining Approach: Exploring Urban Theories on Crime using Mobile Phone Data, *In Proceedings of the 10th International Space Syntax Symposium (SSS 10)*, July 2015.
Chapter 5	The development of a platform to crowdsource safety percetpions towards people at scale	Traunmueller, M., Marshall, P. and Capra, L., streetsmart: Crowdsourcing Safety Perceptions of the lived Urban Environment, *Workshop paper at the 18th Conference on Computer-Supported Cooperative Work and Social Computing (CSCW 2015)*, February 2015.
	The evaluation of our platform	Traunmueller, M., Marshall, P. and Capra, L., Crowdsourcing Safety Perceptions of People: Opportunities and Limitations, *In Proceedings of the 7th International Conference on Social Informatics (SocInfo 2015)*, December 2015.
Chapter 6	The evaluation of a platform to crowdsource safety perceptions towards urban places, including built environment and people inhabiting it	Traunmueller, M., Marshall, P. and Capra, L., "...when you're a Stranger": Evaluating Safety Perceptions of (un)familiar Urban Places, *In Proceedings of the 2nd International Conference on Urban Internet of Things (UrbIoT 2016)*, May 2016.

Table 1.1: Publications of this thesis

1.6 Thesis Structure

This thesis is structured as follows:

Chapter 1 gave an overview on the topic, stated the research problem, formulated the research question and outlined the expected contributions.

Chapter 2 will give an overview on the literature and work that has been done on crime and fear of crime in cities.

Chapter 3 will outline the quantitative research methods that have been used in this thesis, that is *mining big data sets* and *online crowdsourcing*, and their challenges.

Chapter 4 will present our methodology of mining a large, passively collected mobile phone dataset to evaluate urban crime theories at scale.

Chapter 5 will present our methodology of collecting actively perception data following an online crowdsourcing approach, to evaluate at scale safety perceptions towards people.

Chapter 6 will use findings from Chapter 5 and use a similar methodology of online crowdsourcing to explore at scale the effect of presence of people in the urban environment on safety perception.

Chapter 7 will summarize our work, discuss limitations and give future directions.

Chapter 2

Literature

In this chapter, I review theories from urban, social and criminological studies on crime and fear of crime in the city. I then review current research that has been done to develop new methods of evaluation in the domains of crime and fear of crime in cities.

2.1 Crime

2.1.1 Architectural Theories on Urban Crime

Crime in cities has been researched extensively in urban studies and architectural theory (Wood, 1967; Ray, 1971; Brantingham and Brantingham, 1981). Most well known architectural theories about the relationship between *people dynamics*, defined by people's demographic properties such as age, gender and ethnicity, the urban environment and crime lead back to the studies of Jacobs (1961) and Newman (1972).

Due to her studies on what makes a city exuberant or livable in general, Jacobs' research includes a large body of work on urban safety. The term of 'eyes on the street' relies on observations she has conducted on individuals in various American cities, focussing on neighbourhoods in New York City, such as Greenwich Village, her home neighbourhood. The term describes a natural policy mechanism leading to increased safety, based on "natural surveillance" by urban population – or "natural proprietors" of the street – taking "ownership" of a street or an urban place. She suggests that 'eyes on the street' are supported by population diversity (especially in terms of population's age) and a high ratio of visitors to an area, leading to an increased activity at different times. In doing so, she proposes four key elements in urban design, supporting quality of urban environment safety:

- *Mixed use* – Mixed use urban environments allow the usage of an area from different people at different times of day. According to Jacobs, this activates the streets throughout the day and night and supports diversity of population, who provide 'eyes on the street'.

- *Short blocks* – Jacobs suggests that small scale urban design supports urban walking and is therefore beneficial to an area's activity.

- *Diversity in building age and condition* – According to Jacobs, a high building diversity leads to diversity of local residents, from different social background, age and ethnicities.

- *Density* – Jacob suggests that dense urban neighbourhoods are safer than less dense ones, as for instance suburban neighbourhoods, as density supports 'eyes on the street'.

On the opposite, Newman coined the term of 'defensible space', suggesting that neighbourhoods and urban areas should be well defined in terms of their ownership, to become safe. Newman's theory and work on urban design principles is based on studies, focussing on residential areas, in particular public housing projects in South Bronx and Yonkers / New York and Dayton / Ohio, and defines four key concepts:

- *Territoriality* – The concept of *territoriality* suggests that urban space can be designed to affect the perception of ownership. By using physicial design elements, such as fences, gates or signs, urban space can be well defined in what is public, semi–public or privat. In doing so, urban space is subdivided in smaller areas of "ownership", leading to an increased encouragement of residents to take responsibilty for the area. This leads on one hand side to a higher awareness about activities, on the other side such visible markers act discouraging for outside intrudors to commit crime.

- *Surveillance* – In close relation to *territoriality* stands the concept of *surveillance*, suggesting that urban space can be designed for residents to provide visibility about activities on the streets. This includes for instance the definition of sight lines in an area, the location of public spaces or the visibility of elevators. Opposit to Jacobs' theory of 'eyes on the street', where a high visibility is provided through a high activity within an area, this concept focus on the surveillance from local residents towards the streets, and again creates a sense of "ownership".

- *Image / Millieu* – The principles of *image and millieu* relate to public housing and describe the negative affects of large–scale housing developments related to their low–rise environment. Through their exposure, large–scale developments create a stigmatization of residents and are seen as crime attractors for an urban area.

These apparently conflicting theories might result from the methodologies they are based on. Hypotheses from urban design studies, such as the ones stated above, have been critisized for being pseudo–scientific, as long as they use a mixture of personal observations, anecdote and reference to other literature as method. Marshall (2012), for instance, states that researchers still do not know to what degree Jacobs findings are true or not, as there is limited testing. Some researchers attemted to further test these theories, as for instance Weicher (1973), who tested Jane Jacobs results by converting her hypotheses into equations. Findings show little evidence to support her work, except for statements that were already known from common city planning. A few years later, Schmidt (1977) discussed Jacobs theories for the case of Denver, using regression analysis, and also found little evidence to support her work.

Other follow–up studies found more support instead. Fowler (1987) for instance found a relationship between crime and neighbouring to physical diversity, supporting Jacobs work. Felson and Clarke (1998) have proposed the 'Routine Activity Theory', that studies *people dynamics*, such as age and gender, and crime in relation to specific points of interest; they have found that venues such as bars and pubs attract crime by pulling strangers into an area; the presence of middle–aged women on the streets detracts crime instead. Other work discussed the relationship between usage of urban areas, as defined by its zoning, and crime for the case of Los Angeles (Anderson et al., 2013). Therefore researchers conducted an empirical study on the effect of zoning to crime for 205 blocks in eight high–crime neighbourhoods with similar population demographic characteristics, but different zoned land use. Findings suggest that mixed residential and commercial neighbourhoods are safer than commercial–only neighbourhoods. In the same study, researchers also show that neighbourhoods that underwent zoning changes over the years, supporting their diversity, improved their safety through a significant decline in crime, compared to others. Overall, these findings show that mixed–use neighbourhoods, attracting a high diversity of people, are safer than homogenous neighbourhoods.

Based on qualitative studies, Jacobs' and Newman's theories suggest different ways to design the built environment so to take advantage of the resulting social control of crime. But which one applies *where*, and also *when*? How do we know that theories developed in the '60s and '70s are still valid fifty years afterwards? As Marshall (2012) critizises, research still relies on results of these, rather old, works, instead of more recent findings, due to their variance in outcome.

To gain a deeper understanding of the context within which a certain theory holds, researchers need a new perspective on the old social science problem of validating urban crime theories, that scales up in terms of the geographic urban areas under examination, the population sample captured, and the frequency with which studies can be repeated.

2.1.2 Research Methodologies to study Crime

The above outlined theories on urban crime have been evaluated using small scale qualitative studies, such as questionnaires and observations (Jacobs, 1961; Newman, 1972; Felson and Clarke, 1998). Such methods offer very detailled in–depth insights, but are very expensive and time–consuming to run, so studies are usually restricted to a rather small number of people (relative to the overall urban population) and constrained geographic areas (e.g., a neighbourhood); furthermore, they are almost never repeated over time, to observe potential changes. Therefore it becomes very difficult to collect sufficient evidence to understand to what extent a certain theory explains perceptions and behaviours relating to crime.

In recent years, open data movements have made available large repositories of crime data to the public. These circumstances have been useful to start studying crime from a different perspective. Data mining has become a popular method for crime research to detect crime patterns in an urban environment. Recorded crime data has been extensively mined to identify crime hot spots within a city (Paynich, 2013; Wang et al., 2013). A crime hotspot is defined as an area with high criminal activity in immediately surrounding areas. Hot spots provide researchers with a pattern for crime distribution, shape and orientation (Chainey et al., 2002; Eck et al., 2005), and can be even used for crime predictions (Chainey et al., 2008).

For instance, Johnson and Bowers (2004) analysed police recorded crime data for burglaries for the UK county of Merseyside, using statistical techniques that have been used to study disease spreading. Results show a clustered spatial and temporal distribution of crime: as researchers found that burglaries happened in close spatial (300-400m from a prior burglary) and temporal proximity (1-2 months after a prior burglary), they suggest that a preventive action can be taken after a burglary has happened. Milo et al. (2012) use crime data for the same geographical region to define hotspots for parameterized street–level crime. In doing so, researchers found differences in trends for crime activity, as for instance hotspots were found with increasing or unstable crime levels depending on time.

Besides being used in a *retrospective* way to identify crime hotspots in cities, Cheng and Adepejue (2013) use crime data in a space–time scan statistical approach to detect emerging crime patterns *prospectively* at detailled spatial and temporal scale for the London borough of Camden. Instead of following the common method of scanning through all datasets to define 'historical' clusters (such as, all clusters at any time within a defined study period), reseracher detected proactive clusters that started on a specified defined surveillance date. Results were compared to outcome of the traditional retrospective approach, showing the capability of this approach to detect rapidly evolving space–time crime clusters within a spatial area.

While over the years clustering techniques have been refined by improving their clustering algorithm (Adepeju et al., 2015; Wang et al., 2013), they focus on crime density only and hence, do not put crime in any relationship to its environment: they are capable of signaling where crime will happen, without shedding light on the possible *reasons* for that crime. According to Jacobs and Newman, the reasons for crime to happen are to be found in the built environment and the population that inhabits and uses it; different methods are required to quantitatively validate such theories.

Recent architectural and urban design research has attempted to describe the relationship between the built environment and crime. Wolfe and Mennis (2012) discuss the influence of green space in relation to crime by using satellite images to detect green urban spaces and compare them to recorded crime data. Findings show clearly that well maintained green spaces contribute to less crime through an increased community activity and supervision, as also originally suggested by Jacobs. Hillier and Sahbaz (2009) discuss Jacobs' and Newman's theories using detailed spatial data about accessibility of the street network in a London borough, to evaluate correlations with recorded crime numbers. Findings show, for instance, that local movement within an area is beneficial to safety, while global movement from outside into an area is not. Furthermore the study supports the theory that a high mix of use is beneficial to safety.

In a follow–up study (Sahbaz and Hiller, 2007), the same researchers incorporate 'Routine Activities' theories, and explore them using space syntax measurements (Hillier and Hanson, 1984). The work investigates the relationship between street crime occurrences (categorized as 'snatch', 'thread' and 'attack' crime) and the spatial layout of the street network for a London borough. Findings show an overall higher crime distribution along main roads compared to side roads, with the ratios changing throughout the day: the accumulation of 'snatch' crimes increases in the morning and evening hours dramatically, showing that up to 95% of all incidents happen on main roads during these hours.

In a companion paper to work mentioned above (Johnson and Bowers, 2004), Bowers and Johnson (2005) discussed several properties found in urban design principles, as for instance, the location of houses (related to streets and to eachother) and their architecture (such as their floor plan layout) for Merseyside, UK. Findings revealed that for instance, houses at greatest risk are those on the same side of the road, in immediate neighbouring properties and with similarity in their structure of houses where burglaries happened before. Clustering of crime hot spots was found most prominently on straight roads and less on curved roads. In terms of floor plan layout – if same or mirrored to the adjacent house – results showed no significant effects for risks of being burglared.

Davies and Johnson (2015) followed a quantitative network analysis approach to test hypothesis based on the crime pattern theory, suggesting that the configuration and design of road network affects the spatial distribution of burglaries, in particular in terms of street usage, for Birmingham, UK. In doing so, they computed the theoretical metric of betweeness, measuring the level of usage per street segment on shortest paths, and related the level of usage of a street segment to crime. Results support crime pattern theory by finding a higher risk of burglaries on roads with higher usage. Furthermore, results found a lower risk of victimization on straight roads.

These works show that there is a strong relationship between the built environment and location of crime. However, the findings above also point to the fact that there is a third and important dimension to the problem: *people dynamics*, which is described by demographic properties, such as age, gender and ethnicity, of the population. The very same built environment is appropriated and used by different people for different purposes and in different ways throughout the day. *People dynamics* thus need be quantitatively explored in relation to crime too.

When it comes to analysing crime in relation to people, social and criminological research often uses census data. For instance, Tan and Haining (2009) use spatial data of crime and census data to explore the impact of crime on population health for the city of Sheffield, UK. Song and Daqian (2013) explored relationships between spatial patterns of property crime and socio–economic variables of a neighbourhood. Christens and Speer (2005) use census data to explore the relationship between crime and population density, following Jacob's hypothesis that high population density would predict reduced violent crime; they found the hypothesis to be true for densely populated urban areas, but failed in suburban areas where population is less dense.

Census data has also been used in combination with geographical data, amongst other data sources, to describe geographic areas in terms of their *geodemographic* classifications. These classifications provide a summary of an area's demographic, social, economic and built properties, allowing comparisons between them, as for instance comparing different parts of a country. For the UK, two of the most prominent classifications are the Output Area Classification (OAC) at national (ONS, 2011), and the London Output Area Classification (LOAC) at London level (LOAC, 2011; Singleton and Longley, 2015), defining geographic areas into 'Super Groups', 'Groups' and 'Sub Groups', such as 'Intermediate Lifestyle' or 'Ageing City Fringe'. Besides offering opportunities to the public sector to study a variety of urban phenomena (Brown et al., 2000; Longley, 2005), geodemographic data has been used extensively to study crime as well.

Hirschfeld and Bowers (1997) discussed, for instance, the impact of social cohersion, such as the level of 'social control' and 'ethnic heterogenity', in disadvantaged areas on crime levels, and used geodemographic classifications to define disadvantaged areas, besides the 'Index of Local Conditions' – UK's official deprivation measure. Results show that levels of crime were found to be significantly lower in disadvantaged areas with strong social cohersion, compared to others. Furthemore, researchers discussed the effect of Homewatch schemes to the level of burglaries and found such schemes to have a positive effect in affluent, and a negative effect in disadavantaged areas.

Ashby and Longley (2005) uses MOSAIC geodemographic classification (one of the alternatives to OAC and LOAC) to investigate its suitability to define areas for policing purposes at the example of Cornwall and Devon, UK. In doing so, researchers appended geodemographic codes to crime data records for the year 1999-2000 to analyse incidence of crime, offenders and victims relating to 52 MOSAIC neighbourhood types. Results show for instance, that while only 0,4% of the population in the study area live in the 'Council Flats', this type of neighbourhood accounts for a three–times higher crime rate compared to the average. Furthermore, in terms of temporal variation in neighbourhood crime rates, results show for instance, that areas of type 'Country Dwellers' were found to be safer (below the average) than areas of type 'Victorian Low Status' at night.

More recently, Gale et al. (2015) used open source geodemographic data for London (OpenGeodemographics, 2016) and found significant differences for burglary rates in the city, related to the 8 defined geodemographical Super Groups. Results show for instance, that for the Super Group of 'High Density and High Rise Flats' burglary rates were 30% lower, while for 'Settled Asians' they were 25% higher compared to the London average. Closest to the city's burglary average came the groups of 'Intermediate Lifestyles' and 'City Vibe'.

While shedding light into some important relationships between crime and demographics, geodemographic and census data is limited, in that it only offers a static image of the city (i.e., where people reside), without disclosing where people actually spend time throughout the day. Furthermore, census data is only collected every few years, so the information it provides may become quickly stale, especially for areas undergoing massive urbanization processes. According to Jacobs and Newman, it is these *people dynamics* that have great impact on the crime activities of a place which change steadily over time and space, so that we cannot use census data to analyse them.

People dynamics have started to be inferred from geo–located social networks, and used for different purposes. For instance, Prasetyo et al. (2013) use Twitter and Foursquare data to analyse the impact of major natural disasters on people; they do so for haze events in Singapore, and discuss how their approach can help both the private and public sector to better prepare themselves to similar future events. Wakamiya et al. (2013) use geo–located Twitter data to examine crowd interactions, from which social neighbourhood boundaries are defined, thus expanding upon the traditional concept of spatial, administratively–defined neighbourhoods. Discussing crime, Williams et al. (2016) use a set of geo–located tweets for Greater London to mine 'broken windows' indicators (Wilson and Kelling, 1982) of the tweet content and relate them to actual crime data. The work hypothesizes that people, or 'human sensors', use such indicators in their communication, such as on online social media, which can be identified as signal for social decay of urban areas. Results show that mining such signal words have potential to measure social disorder on borough level.

Bendler et al. (2014) explored Twitter patterns related to different crime types. In doing so, they found that crime types, such as anti–social behaviour and homicide, showed differences in tweet patterns through absence already before crime happened and suggest the approach's predictive appeal.

Also discussing tweet patterns related to crime, Kounadi et al. (2015) use a pre–selected data set of tweets containing homicide–related words for Greater London, to explore people's perception of crime in dependence to tweet propensity and spatial proximity to the actual event. The work shows that temporal and spatial proximity matters to the degree of spreading the news on the social network: Temporally, results show that more than half of all homicide related tweets happened in the first week, and overall the majority of tweets happened within one month after the crime happened. Spatially, results show that user's proximity of home location matters to if a crime was tweeted about, or not. Futhermore, they found that certain characteristics related to the crime (as for instance, if the victim was young, the presence of a knife or if it was gang related) impact posting frequency.

Malleson and Andresen (2014) used Twitter data to investigate people's risk of becoming a vicitim of violent crime at the example of Leeds, UK. With this work, researchers critisize the use of outdated and static census data and suggest tweets as proxies for mobile population that can be used to model crime risk following various spatial analysis approaches. Results suggest alternative hotspots for violent crime outside the city centre, which have not been captured using conventional approaches.

Besides it's analytic value, Twitter data has also been used to predict crime. Wang et al. (2012), for instance, use sentiment analysis to relate content of Twitter messages to hit–and–run crime activity and show opportunities to use this approach to predict crime. Gerber (2014) identifies automatically discussion topics of tweets for the city of Chicago, using statistical topic modeling, and adds them to a standard crime prediction model with the aim to improve its performance. Results show that by including these topics, the model increased its performance for 19 out of 25 discussed crime types, based on kernel density estimation, with highest increase for the types of 'Stalking', 'Criminal Damage' and 'Gambling'.

These example show that social media, in particular Twitter, is a rich data source from which to derive information about *people dynamics*; however, it is also unrepresentative of the whole urban population, because of high bias in its adoption (Boyd and Crawford, 2012). An alternative data source that can be used to mine *people dynamics* in urban areas, and that is subject to significantly lower bias than social media, is telecommunication data.

Telecommunication data has been recently used to understand the relationship between cities (and even whole countries) and socio–economic deprivation, both in the developed world (Eagle and Macy, 2010) and in developing countries (Smith Clarke et al., 2014). In relation to crime, recent work (Bogomolov et al., 2014) uses a mobile phone data set to extract human behavioural data and combines it with census data, describing London's population per borough. The two datasets were then related to crime data with the aim of predicting crime hot spots for urban areas of London. By including human behavioural data to the census data, researchers were able to explain up to 70% of crimes, showing the importance of including dynamic properties, as derived from mobile phone data. Furthermore, the study supports the importance of people diversity in relation to urban safety, suggesting that a high diversity of people leads to less crime, as described by Jacobs (1961).

Malleson and Andresen (2016) use, besides census and Twitter data, a similar mobile phone data set as used in (Bogomolov et al., 2014) to evaluate the most appropriate measure among them for ambient population–at–risk for the crime type 'theft from a person'. Results suggest that census data for workday population is the most suitable predictor from discussed sources. These results have to be taken with care though, since they only apply to the case of London, leaving questions unanswered about the method's performance in another city or another culture.

All the above works show the potential of data coming from various sources to research crime in cities. However, focussing on prediction, none of them has been used in a descriptive way, such as to evaluate the established theories that current urban design principles are based on. I believe the same data can be used to understand established theories as well. In doing so, findings can be used to decide how to design cities that will incur in less (fear of) crime; in Chapter 4 I will illustrate how.

2.2 Fear of Crime

2.2.1 From Crime to Fear of Crime

In the previous section I have outlined theories and recent research on crime in the urban environment. However, looking at the situation of urban development nowadays, we observe that besides crime, fear of crime has become an increasing problem for the broad population that has to be included into the discussion. Fear of crime describes a perception, such as a lack in feeling of safety in an environment. According to Doran and Burgess (2012) there is a strong relationship between crime and the fear of crime in an urban environment, as there is a higher fear of crime among the population of areas with high crime rates. However, other researchers (Matei et al., 2001) suggest that fear of crime is not limited only to victims of actual crime, but even more affects the broad population supported by modern media, such as online social networks and television broadcasting. In doing so, fear of crime is not restricted to space and time and has the potential to be more widespread than crime; hence, it is seen as an even bigger problem than crime itself (Brown and Polk, 1996; Oc and Tiesdell, 1997). Doran and Burgess (2012) describe the link between real crime and fear of crime as 'vicious circle', *"because it results in residents adopting protective and avoidance behaviours which contribute to the breakdown of informal social control, more fear of crime and crime itself"*. In this sense, fear of crime limits a city's population in their daily actions by avoiding public space that is perceived to be unsafe (Nasar and Fisher, 1993; Warr and Ellison, 2000), leading to reduced quality of urban life (Pacione, 2003). For the city itself, this results in a problem for its sustainability, as avoiding public space leads to a decrease in the city's walkability and increased motor traffic (Warr and Ellison, 2000).

Fear of crime manifests through demographic, social and environmental measures of the city and its inhabitants that research aims to describe through a number of established hypotheses. We will briefly outline these hypotheses before discussing the research methods they derived from and current work.

2.2.2 Theories on Fear of Crime

Research suggests three main groups of hypotheses relating people's fear of crime to both dynamic (social, demographic) and static (environmental) properties of a city.

- *Demographic–based* hypotheses aim to explain differences in fear of crime perception and feeling of vulnerability associated with different demographic groups, in particular in terms of age (Zako, 2009), gender (Felson and Clarke, 1998) and ethnicity (Day, 1999; Pain, 2001).

- *Social–based* hypotheses suggest that fear of crime is the result of a general state of anxiety, caused by social disorganization resulting from social change (Furstenber, 1971), subcultural diversity (Merry, 1981) and community concern (Lane and Meeker, 2003). Beck (1992) describes this state of anxiety as 'risk society', in which *"fear of crime is conceptualized as an expression of people's wider feelings of insecurity or uncertainty about life"*. This insecurity leads to a state where the 'unknown', as people on the streets, are perceived as dangerous, resulting in avoidance behavior in every–day life (Lianos and Douglas, 2000).

- *Environmental–based* hypotheses relate fear of crime to the static built environment, suggesting that fear of crime is a result of how people experience and interpret urban space (Bannister and Fyfe, 2001).

Based on the above established hypotheses, we can summarize that fear of crime is a result of *dynamic* and *static* factors found in the city, described as separate entities in various established hypotheses. *Dynamic* factors describe demographic and social properties of people, that have impact on different fear of crime perception among the urban population. *Static* properties describe variables found in the built environment and include both environmental cues of social disorder and cues resulting of urban planning decisions or the lack thereof, as for instance the lack of 'natural surveillance'.

The urban environment constitutes a composition of *places*, defined not only by its physical *or* human aspects, but by *both* interrelating (Tuan, 2001). Human experiences and senses shape these places and add meaning to them, including the perception of safety. Hence, to understand fear of crime in an urban environment it is necessary to include both *static* and *dynamic* properties into the methodology of research informing each other.

Next I will outline fear of crime research methods these hypotheses are evaluated by and review state–of–the–art work that has been done in the field.

2.2.3 Research Methodologies for Fear of Crime

Qualitative studies. Throughout social and criminological research, the most common used method to research fear of crime is through victimization surveys, public perception questionnaires and semi–structured interviews, which are mostly conducted at home, detached from urban space (Fountain, 2012). Using cognitive mapping and collective fear mapping methods, spatial and temporal effects are being investigated by activity diaries that are being completed at a separate location other than the one being depicted. Because of this reason, they have been criticized for delivering a rather general image of safety perception (Farrall et al., 1997; Jackson, 2005). Studies discussed for instance the phenomenon of the *Familiar Stranger* (Milgram, 1977; Paulos and Goodman, 2004), suggesting that familiarity with a situation is a key element to our safety perception: the more familiar we are with our surrounding, the safer we feel (Nasar, 1994). Other work used these methods do develop theories linking different properties of the city to resulting perceptions of the population. Some theories have explored the relationship between safety perception and the built environment and found that environmental cues, such as broken windows (Wilson and Kelling, 1982), graffiti, abandoned buildings and broken streetlights (Doeksen, 1997; Ross and Mirowsky, 1999; Skogan, 1999) can trigger a feeling of unsafety through being perceived as warning signs for crime activity in an area (Mirrlees-Black and Allen, 1998; Tulloch, 2000). Other cues can be a result of urban planning, or the lack thereof (Gehl, 2010). Such environments are perceived as unsafe because they appear to be attractive sites for criminal activity, supported by the lack of 'natural surveillance' (Jacobs, 1961).

Other theories still have explored the relationship between perceptions of safety and people, both in terms of *who we are*, and *who we see*. People perceive safety differently depending on demographic properties, such as age (Zako, 2009), gender (Felson and Clarke, 1998) and ethnicity (Day, 1999; Pain, 2001). For instance, research has found that the most fearful groups are women and the elderly, who are surprisingly least at risk of being victimized (Katz and Webb, 2003; Painter, 1996; Pantazis, 2000; Taylor and Hale, 1986), whereas young men, who are most at risk, show the least fear of crime (Warr, 1984; Hollway and Jefferson, 1997).

Following a similar, more graphical approach stemming from the urban geography, urban designers and geographers use mental maps to describe urban space and its differences in fear perception of the inhabitants. Mental maps allow researchers a very personal 'image of the city' (Lynch, 1960) relying on memories and experiences of study participants through simple freehand sketches and notes relating to a geographic area. Matei et al. (2001), for instance, use mental maps to explore the role of the media in relation to people's fear perception in Los Angeles. The study uses Geographical Information System (GIS) technology to process hand–drawn mental maps taken in seven neighbourhoods all over the city by 215 study participants. Analysis revealed that the concentration of certain ethnicities has a more significant impact on participant's fear perception than actual crime activity. Another study (Hallman et al., 2013) used mental maps to explore relationships between usage of public space and the violence perceived by adolescents in South Africa, differentiated by gender, age and residential background (urban – rural). Findings suggested that teenage girls show most significant movement restrictions in public areas, compared to the other participant groups, indicating that this group perceives more dangers in their communities than others.

These methodologies offer researchers a very rich and detailed insight into what triggers people's fear in urban space. However, relying on memories and experiences of the participants, they do not take situational factors into account that might affect the outcome. Miller (2008) found that factors such as the visual appearance of the built environment and the people inhabiting it have an impact on the outcome of fear of crime surveys when taken in situ. To bring such situational factors from the wild into the laboratory environment, recent work uses virtual environments (VE) incorporating 360 degree images. For instance, Park (2008) developed a VE pedestrian model representing a fear generating area and discussed fear of crime impact on routing behavior of people in the city. As a follow–up (Park et al., 2011), the same author used a VE in a study focusing on elderly people, for whom common methods of fear of crime measurements, such as surveys and questionnaires, were found to be too complicated to use. Findings show a clear improvement in terms of accuracy and depth of outcome by using VEs and point out opportunities to support traditional fear of crime research methods. Cozens et al. (2004) used a VE to discuss fear of crime perception on British railway stations. The work uses images to create Quick Time Virtual Reality (QTVR) walkthrough scenes and found that visibility at stations and the design of the station shelter is crucial to passenger's fear perception.

These qualitative methodologies offer a semantically rich insight into people's safety perception and offer opportunities to *develop* theories. However, being expensive to conduct and very time and resource consuming, they show limitations in scale and replicability. It is difficult to reproduce such studies with a larger number of people and over different time scales and spatial extents to investigate possible changes relating to temporal and spatial differences. These circumstances make it difficult to use such methodologies to *evaluate* theories of fear perception in the urban environment at scale.

Quantitative studies. Technology developments over the last years allowed researchers to study perceptions, such as on safety, quantitatively at a large scale, using mobile technology and online crowdsourcing approaches. With the rise of mobile devices such as the smartphone, mobile technology offers new opportunities to study safety perceptions in urban space quantitatively, and in–situ. Researchers use this technology, for instance, to collect data describing people's body feedback. For instance, Nold (2009) used GPS and sensors on study participants' skin to measure galvanic skin response while walking through the city. Resulting data was mapped, showing different levels of arousal of the participant depending on environmental changes, such as the presence of cars or other people. More recent work measures human brainwaves using electroencephalography (EEG) (Mavros et al., 2012, 2016) while walking through urban environment. Results show how different urban environments, such as a noisy main road compared to a side road, affect people's brainwave activity, and how EEG technology can be used to relate this activity to the urban environment. The recent advent of wearable technology, such as Fitbit (2012) or smartwatches (Apple, 2015) supports such approaches to measure body response data by bringing it on a broad scale.

Besides measuring body feedback by means of passive sensors to map perceptions, developers created also a variety of mobile applications to actively collect safety perceptions, to record and map crime and fear of crime data. A common concept for such data gathering applications is the "Panic button" design, that aims to provide the user with a higher feeling of safety through the conciousness of not being alone in the urban environment (RedPanicButton, 2014; EyeOnMe, 2014), in a Taxi (Taxiaviso, 2014), at school (SchoolGuard, 2014) or while taking part in a demonstration (PanicButton, 2014). This is accomplished through a steady connection on social media channels, email and text messages to friends, family or authorities. Other concepts aim to increase the user's safety perception through total surveillance of the immediate surrounding through technology implemented in the device, such as the camera and microphone (iWitness, 2014). However, as the main purpose of such applications is to increase the feeling of safety of the user in a situation perceived as unsafe, the gathered data offers only limited research purposes: Firstly, collected data covers geographical and temporal information through application usage only, which is the location and time–stamp of unsafe perceived situations. Secondly, the data does not provide further details about any reasoning for the user's perception in that situation, and hence keeps researchers in the dark about any background information.

Recent work in social computing suggests applications to crowdsource such detailed information on a current situation. For instance, *VoiceYourView* (Lam et al., 2011) is a research project that discusses various crowdsourcing applications, both for the mobile phone and for publicly deployed urban kiosks, with the aim, to encourage people to share their opinions towards the environment to make them feel more inclusive. In doing so, researchers collected 2000 design critiques on public space design of 600 users by an intelligent kiosk deployed in public space, that allowed unstructured voice and text input (Whittle et al., 2010). Using natural language processing and speech recognition (Simm et al., 2010; Nasa et al., 2010), critiques were processed accordingly to extract meaning and summaries presented on a public display to engage public conversation. Results show a high accuracy (78%) for auto–summarized comments left by the public, indicating the approach's potential to create a reasonable picture of the public opinion.

Other approaches use applications installed on mobile devices, such as the mobile phone, to measure perceptions of space and place. *Mappiness* (Mappiness, 2013), for instance, is a mobile application that allows its user to record his/her wellbeing in random situations throughout the day, and reflect on them afterwards. At the same time, crowdsourced data can be used for research purposes. In doing so, researchers related for instance people's wellbeing to paid work (MacKerron and Bryson, 2013) and found, that wellbeing varies depending on the location, wether working from home, in the office or elsewhere. Other work (MacKerron and Mourato, 2013) uses mappiness data to support the general assumption, that people are significantly happier outdoors in nature, rather than in the urban environment.

Focussing on safety perception, other work uses mobile applications to collect public opinions of urban places. Applications such as *FOCA* (Solymosi et al., 2015) or *uSafe* (Christin et al., 2013) for instance, allow the user to vote about his/her safety perception for the current location and to reason it, depending on features in the environment. Based on gathered data, maps are being created showing colour–coded results as overlay on the urban street layout so that areas perceived unsafe can be avoided by the user. Similar applications (Walkonomics, 2014) use such data to recommend 'safe' routes to the user with the aim to support a city's walkability. As fear of crime perception is highly subjective, the main limitation such approaches suffer from is their lack of human properties of the application user, such as age, gender and ethnicity which, based on the literature, we identified to have great impact on the voting. In fact, a number of mobile applications have been criticized as supporting prejudice and racism by labeling public urban space based on personal opinions (SketchFactor, 2014; GhettoTracker, 2013).

Reviewed approaches focus on the urban environment without knowing *who* currently inhabits it and therefore exclude *people dynamics* that have, as identified above, great impact on safety perception in an urban place. Furthermore, data collection based on mobile phone applications is very time–consuming and excludes specific groups of people not using smartphones, such as the poor and elders, who are found to experience most fear of crime (Pain, 2001).

To gather large amount of data in less time that includes these groups, research has developed more accessible crowdsourcing approaches via desktop–based applications (Pew, 2014). Recent work suggests online games that are not used in–situ, but enable researchers to collect large amount of perception data in less time: Using photographic images of urban environments from Google Street View (Google, 2016), these approaches aim to recall in–situ experience that focus on the visual aspects of the city. For instance, *Urbanopticon* (Quercia et al., 2013) presents 360 degree 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 (Lynch, 1960), a collective mental map is drawn with the aim of detecting happy places in the city.

Similar approaches have been used to crowdsource other perception data beyond happiness too. For instance, *Urbangems* (Quercia et al., 2014) crowdsources perception data besides happiness about beauty and calmness of a city based on visual cues (Wilson and Kelling, 1982) found on Google Street View images. By showing two random images of Greater London to the user, who is asked to chose the happier, more beautiful or calmer of them, *Urbangems* aims to identify visual cues of the built environment and the perceived attributes people attach to them. In doing so, researchers visually analysed the images in terms of presence of colours, different textures (vertical and horizontal shapes found in the image) and "visual words" (such as interest points in an image that typically correspond to a change in the shown surfaces, e.g. an edge) in relation to user opinions for happy, beautiful or calm rated images. Findings revealed, for instance, that green colours and vertical shapes contribute to the beauty and calmness in urban appearance, while darker shades of colours (e.g., grey, brown or dark red) and horizontal shapes do not.

Following the same methodology, *PlacePulse* 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 American and European cities, such as New York City in the U.S. and Salzburg in Austria (Salesses et al., 2013). Results show two main differences, as perceptions in American cities were found to be more clustered, indicating a high visual variance between urban areas, compared to a rather even distribution for European cities. Furthermore, the research explored the relationship between perceived safety and actual crime activity using the example of New York City and found significant correlations between them: areas with a higher rate of 'class' and 'uniqueness' were perceived as safer and showed a lower crime rate than others.

As a follow–up based on these findings, researchers developed *Streetscore*, an algorithm that identifies visual cues in Google Street View images using computer vision technology to automate the process (Naik et al., 2014). 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.

The use of artificial renderings, such as 2D photographic images, to represent 3D places has been criticized in the past in terms of their representativeness (Rose, 2007); nonetheless, above works show the potential of using online crowdsourcing to gather perception data about safety at scale focusing on visual cues of an urban environment. However, recalling Tuan (2001), we have identified that an urban environment constitutes a *place* shaped by both physical *and* dynamic factors including built environment *and* the urban population. By using Google Street View images, reviewed studies focus primarily on the built environment as this source barely shows people in the images. As Google Street View images are mostly captured in the early mornings and hence show empty sidewalks, little traffic, closed shops and generally little activity, they keep out *people dynamics* which have great impact on the perception of a place (Gehl, 2010; Tuan, 2001). People change the appearance of urban space over time and give it different meaning, while using steady images of static facades only show parts of the visual urban experience. It is thus unclear whether a similar methodology could be used to validate theories about *places*, consisting of built environment *and* people, in relation to fear of crime, and with what results.

In summary I have identified that common fear of crime research using qualitative methodologies is limited in scale and does not deliver answers about fear of crime in relation to in–situ experiences. Mobile applications bring fear of crime research to situations in the wild and on a quantitative level, but are time–consuming in data generation. Furthermore, they do not include information on people dynamics and leave out certain user demographics. Desktop–based crowdsourcing approaches show potential to offer quick data generation and are more inclusive of often under–represented user groups. They have also shown to be acceptable alternatives to in–situ data collection approaches. However, by focusing on the built environment only, they ignore the impact of people dynamics, which literature has shown to have great impact on our safety perception. To include people dynamics, I suggest a similar methodology that builds on reviewed work focusing on the built environment, but extends it by adding the evaluation of 'visual cues' from the urban population to it. Therefore I will develop an online platform to crowdsource safety perceptions towards the appearance of other people using images in Chapter 5, and will relate our findings to the appearance of the built environment in Chapter 6.

In the next chapter, I will outline my new mixed method approach that uses telecommunication data to evaluate urban crime theories, and free available online images to explore the role of *people dynamics* in relation to fear of crime in the city at scale.

Chapter 3

Methodology

This thesis proposes a mixed method approach that uses both passively and actively collected data to quantitatively evaluate crime and fear of crime theories at temporal and spatial scale. The method has been applied to the case of Greater London, UK. London represents a large and complex metropolitan city, composed of many different neighbourhoods, each with its own distinguishing characteristics in terms of built environment and *people dynamics*. It thus represents a case where qualitative approaches to investigate urban crime and fear of crime theories would not scale, both because of the geographic span of the areas to study, and because of the time frequency with which one may wish to repeat these studies (e.g., to observe changes in relation to ongoing immigration processes (Snyder, 2007)). In this chapter I will discuss my approach.

3.1 Passively collected data

With the rise of digital technology over the last decade, the amount of data that has been available has increased tremendously (SINTEF, 2013). Generated passively by its users or from sensors in the environment, and supported by the open–source movement, it offers industry and science in various fields new ways to study human and environmental phenomena at a large scale. Big data sets have been used to *analyse* complex behaviour patterns, and to build *prediction models*, using data mining methods and machine learning algorithms processing the data.

The method I propose in Chapter 4 to evaluate urban crime theories requires access to two types of datasets: one providing information about crimes, and one with information about *people dynamics*. The first is open–source in the UK and made available by two authorities: the Metropolitan Police and the City of London Police (available for download at (PoliceUK, 2013a)). The dataset provides temporal and spatial information about recorded crimes. For people dynamics, I use anonymised and aggregated data collected and made available by a mobile telecommunication provider in the context of a data mining challenge (Telefonica, 2014). Such data provides temporally and spatially detailed information about people dynamics, in terms of how many people of a certain age, gender and type (residents, workers, visitors) are present in a certain area at a given time. Sensitive data such as telecommunication data is usually more confidential and less easy to access, compared to open–source crime data. However, with the ongoing open–data movement and a variety of data mining challenges, such as the Data for Development challenge (Orange, 2013) and the Big Data Challenge (TelecomItalia, 2012), there is a clear trend towards mobile phone providers making their data available to the public.

With the aim of evaluating urban crime theories, I use the data in primarily a descriptive, rather than predictive way. I will first extract metrics from these data to cepture theories; I will then use correlation and regression analysis to describe the relationship between these continous predictor variables and urban crime (Field et al., 2012).

Challenges. Using a passively collected data set has the benefit to validate urban crime theories on a large scale and, if available, to re–run the study for different geographic areas and at different times. However, as research following this approach is dependent on the quality of data that is being provided, methodologies such as the one presented in this thesis, are limited by availability and level of granularity of the data. Therefore, results of such work need to be interpreted with care, as they might suffer of potential flaws that are based on poor data quality of unreliable data. In practise, I experienced such drawbacks, based on limited spatial and temporal availability of my data at hand. In Chapter 4, I will discuss these drawbacks and will show to what extent we can use telecommunication data to evaluate urban crime theories at scale.

3.2 Actively collected data

While presence of people in a certain area / at a certain time can be passively collected simply listening to the signal of the mobile phone we carry in our pockets, perception data is not that ready available. To enable researchers to perform quantitative studies on perceptions, data needs to be generated first. A common way to actively generate such data is *online crowdsourcing*, where a task is being outsourced to the crowd, enabling reserachers to collect data for a specific purpose. Online marketplaces, such as Amazon Mechanical Turk (AMT, 2010), Crowdflower (2014), or Clickworker (2014) commercialize crowdsourcing, enabling an online community of crowdworkers to share and execute such crowdsourcing tasks, making the process very time–efficient.

Besides enabling researchers to quickly collect data to answer specific questions, online crowdsourcing also offers opportunities to gather less rigid *"soft data"* (Cambridge, 2000) that has been difficult to collect quantitatively. Soft data describes, for instance, perceptions or feelings of people in specific situations, which have been captured in the past mostly following qualtitative research methods. As described in Chapter 2, there have been various approaches to obtain people's perceptions of the urban built evironment quantitatively and over a short time, using Google Street View images (Salesses et al., 2013; Quercia et al., 2014).

The method I propose to evaluate fear of crime theories at scale uses online crowdsourcing to quickly collect safety perception data. In particluar, I use online images of people (Chapter 5) and combine them with Google Street View images (Chapter 6). Similar to Salesses et al. (2013) and Quercia et al. (2014), I built two online crowdsourcing platforms called *Streetsmart* (Chapter 5) and *Streetwise* (Chapter 6), presenting images to participants, who were asked to rate and comment on them in terms of safety perception. From the collected data, I extracted metrics representing theories of safety in urban environments, as defined in the literature.

To draw relationships between continuous outcome (perception of safety) based on several categorical predictor variables (*people dynamics*), I use Analysis of Variance (ANOVA) and Analysis of Covariance (ANCOVA) with planned contrasts on collected data (Field et al., 2012). To reach statistical significancy a large amount of data needs to be collected from a large number of participants. For this pupose I recruite crowdworkers from social media (Chapter 5) and from AMT (Chapter 6).

Challenges. Besides the approach's benefits of gathering a large amount of soft data in a short time, compared to qualitative approaches, it is important to be aware of possible drawbacks of the suggested methodology. As research using actively collected data from a crowdsourcing platform is dependent on the the people the data has been collected from, this approach is limited by participant's demographic background. For instance, crowdsourcing has shown to suffer from self–selection bias due to lack of crowd–control when using Open Street Map (OSM) (Quattrone et al., 2015). In practise, I experienced similar challenges in reaching people from a broad demographic variety: As most of my study participants were found to be Caucasian and middle–aged, I am not able to generalize my findings for other demographics and age groups. Furthermore, as a crowdsourcing study is dependent on the study design, there is the potential of being biased by the researcher's background. In my work, when using images representing different demographic groups of people, a broader sample size was suggested to minimize this effect. Still this bias might lead to flaws in results. In Chapter 5 and Chapter 6, I will show to what extent we can use crowdsourced data to evaluate fear of crime theories at scale.

Chapter 4

A Data Mining Approach to Evaluate Crime Theories in an Urban Environment

Part of the work presented in this chapter also appeared as a full paper accepted to the 6th International Conference on Social Informatics (SocInfo) 2014. [Acceptance Rate: 23%]:

Traunmueller, M., Quattrone, G. and Capra, L., Mining Mobile Phone Data to Investigate Urban Crime Theories at Scale, *In Proceedings of the 6th International Conference on Social Informatics (SocInfo)*, November 2014.

4.1 Introduction

The relationship between *people dynamics* and crime in urban environments has been researched extensively in architectural and urban studies over the last decades, with theories that sometimes appear to conflict with each other. Most influential theories lead back to the 1960's and 1970's: Jacobs (1961) argues that population diversity (especially in terms of their age), a high ratio of visitors and workers to an area are supportive to street activity, leading to less crime, due to providing 'natural surveillance' or 'eyes on the street'. On the other side, Newman (1972) hypothesizes the opposite, that diversity actually brings crime to an urban area through providing anonymity among population. He supports a clear separation of public, semi–public and private areas, and states that a high ratio of residents is supportive towards urban safety. In addition, Felson and Clarke (1998) argue that a high ratio of male and young population brings crime to an area.

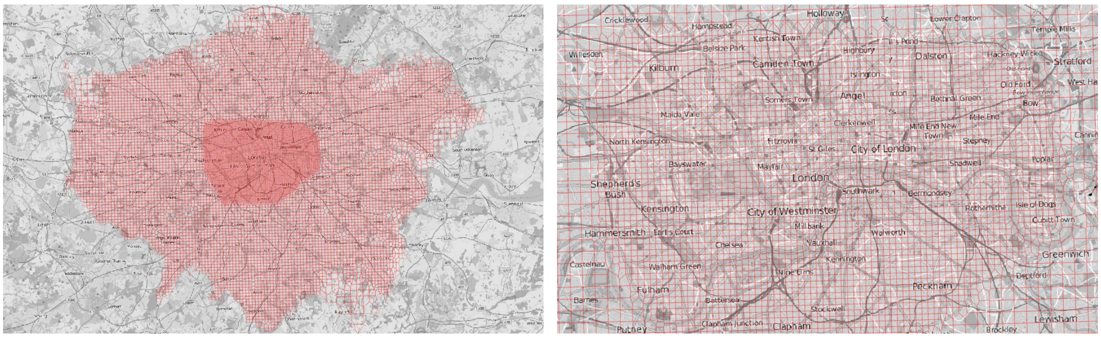
Each theory has been evaluated, and indeed supported, by means of qualitative research methods that enable in–depth investigations into the reasons behind certain phenomena. However, such methods are very expensive and time–consuming to run, so that studies are usually restricted to a rather small number of people (relative to the overall urban population) and constrained geographic areas (e.g., a neighbourhood); furthermore, they are almost never repeated over time, to observe potential changes. It becomes thus very difficult to collect sufficient evidence to explain under what conditions a certain theory holds.

In this chapter I propose a method to quantitatively investigate such urban crime theories at scale, using crime data records and anonymised mobile telecommunication data. The first is open–source in the UK and made available to download by two authorities: the Metropolitan Police and the City of London Police (PoliceUK, 2013a). For the latter I use anonymised and aggregated data collected that has been made available by a mobile telecommunication provider in the context of a data mining challenge (Telefonica, 2014). Telecommunication data is usually hard to access due to confidentiality matters, but data mining challenges, such as the Data for Development challenge (Orange, 2013) and the Big Data Challenge (TelecomItalia, 2012) show a clear trend of mobile phone providers towards making their data available to the public.

From the crime dataset, I extract quantitative information about crime activity, as it happens across different urban areas at a very fine spatial granularity. From the telecommunication dataset, I extract metrics that act as proxies for previously developed urban crime theories that link presence of different people in an area with crime. I can do so as mobile telecommunication data provides a demographic breakdown (by age, gender and type – residents, workers or visitors) of how many people are present in a given area at a given time. As the penetration of mobile phones in cities of developed countries is very high, and as mobile phones are personal devices usually carried by people all the time, I expect such data (and the derived metrics) to offer a rather accurate and fine–grained image of the urban area under examination.

I then follow a statistical analysis approach, using correlation and regression analysis between crime data and the defined metrics to test urban crime theories at scale.

I apply this method to data obtained for the city of London, UK, and find that, in



(a) Greater London **(b)** Inner London

Figure 4.1: Maps showing different cell sizes of used grid. Left map shows Greater London, right map shows Inner London.

this city and at the present time, Jacobs' theory of 'natural surveillance' (Jacobs, 1961) holds: I discover that age diversity, as well as the ratio of visitors in a given area, are significantly and negatively correlated with crime activities; furthermore, Felson and Clarke (1998)'s theory that links a higher presence of young people to higher crime is also confirmed.

The remainder of the chapter is structured as follows: next I present my method, in terms of the datasets I leverage, the pre–processing and data manipulation I have conducted, and the metrics I have extracted as proxies for urban crime theories. I discuss the results obtained when applying my method to data for the city of London, UK, and finally conclude by discussing implications and limitations.

4.2 Method

In this section, I describe the method I propose to quantitatively explore previous theories of urban crime. I start with a brief description of the datasets and identify potential limitations of this approach, based on the dataset quality and availability. I then present the pre–processing steps these datasets underwent, and finally elaborate on the metrics I extracted from them as proxies for urban crime theories.

4.2.1 Dataset Description

The method I propose requires access to two types of datasets: one providing information about people dynamics, and one with information about crimes. The former includes information about people's demographic properties, such as their age and gender, at a certain time and location; the latter contains information about crime activity

Date	*Time*	*Grid ID*	*Total*	*Home*	*Work*	*Visit*	*Male*	*Female*	*0–20*	*21–30*	*31–40*	*41–50*	*51–60*	*60+*
10/12/2012	9:00:00	1122...	430	110	290	30	240	190	0	80	90	120	100	40
10/12/2012	10:00:00	2412...	910	210	160	540	520	390	0	180	180	260	170	120
10/12/2012	11:00:00	1092...	900	570	250	80	520	380	10	160	190	250	210	80
10/12/2012	12:00:00	2124...	690	80	120	490	410	280	10	120	150	190	140	80

Table 4.1: Record sample of mobile phone data, showing the number of people per area, per hour

Crime ID	*Month*	*Reported by*	*Lon*	*Lat*	*Location*	*LSOA Code*	*Crime Type*
df0c4...	2012-12	Met Police	-0.219	51.568	near Clitterhouse Rd	E010...	Burglary
0f9a5...	2012-12	Met Police	-0.217	51.565	near Caney Mews	E010...	Burglary
62235...	2012-12	CoL Police	-0.221	51.570	near Claremont Way	E010...	Crim. damage & arson
194ed...	2012-12	CoL Police	-0.222	51.563	near Petrol Stn	E010...	Crim. damage & arson

Table 4.2: Record sample of open crime data, showing crime incidents, geo location and crime type

(again at a certain time and location). For this study, I chose datasets that cover the city of Greater London, UK. I did so as London represents a large and complex metropolitan city, composed of many different neighbourhoods, each with its own distinguishing characteristics in terms of built environment, demographics, and people dynamics. It thus represents a case where qualitative approaches to investigate urban crime theories would not scale, both because of the geographic span of the areas to study, and because of the time frequency with which one may wish to repeat these studies (e.g., to observe changes in relation to ongoing immigration processes (Snyder, 2007)).

People dynamics. I use anonymised and aggregated data collected and made available by a mobile telecommunication provider in context of a data mining challenge with a 25% penetration in the UK. The dataset contains 12,150,116 footfall count entries for the Metropolitan Area of London for the course of 3 weeks in December 2012/January 2013. The geographic area is divided by the data provider itself into 23,164 grid cells of varying size, as shown in Figure 4.1: for the more densely populated areas within inner London, the grid size is about by 210×210 meters, while for the less dense areas of Greater London, the grid size increases to about 425×425 meters. For each cell, footfall counts are given on a per hour basis over the three week period, further broken down by gender (number of males/females), by type (number of residents, workers, visitors) and by age group. Table 4.1 shows a sample of my mobile phone dataset.

Crime data. I use open crime data records (PoliceUK, 2013a), which, for the area of Greater London, are made available by two authorities: the Metropolitan Police and the City of London Police. These records provide information about the reporting police district, the exact location (longitude and latitude) of the crime, the name and area code of the crime, and the crime type (which the UK police differentiates into 10 categories: anti–social behaviour, criminal damage and arson, other theft, other crime, violent crime, vehicle crime, burglary, shoplifting, drugs, robbery). Unfortunately, no timestamp is given for when the crime took place/was reported, and the only temporal information I have is the month during which it took place. I thus collected crime data for the months of December 2012 and January 2013 (to temporally match my mobile phone data), and retrieved 83,526 recorded crimes in total. Table 4.2 shows a sample of my crime data set.

4.2.2 The challenges of using passively collected data for studying urban crime

In proposing methods that aim to validate theories based on passively–collected data, such as mobile telco data, one has to assess the quality of such data to begin with as findings will depend on it. Looking at the two datasets my proposed method relies on, I already can identify a number of potential limitations that will affect our findings.

Mobile phone data. The mobile phone data set has a number of flaws related to reliability and representativeness that need to be pointed out, to enable a correct interpretation of the results. First of all, having the devices, and hence their users, geographically defined by a grid, rather than the exact geo–location, introduces a problem of geographical accuracy. Especially in an organically grown city such as London, where streets and boroughs change dramatically within a close range, this becomes a problem when relating findings to urban areas. Furthermore, these circumstances make it difficult to differentiate between people on the streets or in buildings.

Second, as can be seen in Figure 4.3(b), mobile telco data shows a high concentration of visitors at the far outer skirts of Greater London, which seems rather surprising. This is an area partly covered by a major road (M25) with high traffic. It sits right at the boundary of the area for which telco data was provided. This might signal data unreliability.

Another point worth mentioning is that, even after discussions with the mobile phone provider, it stays unclear how different types of people (such as, residents, visitors and workers) have been identified, same as how cleansing of the data was performed, as such information is confidential. This uncertainty leads to flaws that might affect my findings and I need to be aware of.

Crime data. Looking at the data quality of the open source crime data set, I can identify a number of limitations that will affect the study's results. The data set includes a geo–location of each data point. However, according to the open source platform these coordinates only represent an approximate and not the exact location of a crime, defined by a master list of over 750.000 'anonymous' map point (PoliceUK, 2013b). This results in issues for the study related to accurately pinpointing crime location. Furthermore, my results will only refer to reported crime as the data set provides information about reported crime, even though it is well known that often crime is not being reported (HMIC, 2014). The fact that each data point provides temporal information of when a crime has being reported (per month) leads to issues of identifying when a crime actually took place. As reporting times for different crime types may differ (HMIC, 2014), this will affect the validity of the results. In the worst case, this means that some crime types cannot be considered (e.g., if a crime type typically is reported only months after it occurs).

The majority of recorded crimes are geocoded by police forces using a variety of methods such as tagging by mobile GPS receivers or address referencing. However, such data in raw form present a high risk to individual disclosure (Kounadi et al., 2014), and as such, it is not possible to publicly release crime data of this level of precision within a UK context given legislative constraints (Singleton and Brunsdon, 2014). As such, publically accessible crime data released through PoliceUK (2013b) are anonymised so that no individual crime event location is identifiable (Tompson et al., 2014). Crimes are allocated to a nearest centroid point of a pre-defined zonal geography (Tompson et al., 2014) which represent a collection of streets. This geography was created using Voronoi polygons drawn around street segment centroids and points of local relevance. To ensure privacy, polygons were merged if necessary to ensure each contained at least eight addresses (Singleton and Brunsdon, 2014; Tompson et al., 2014). Data made available by PoliceUK (2013b) uses these centroids as the recorded location of any crimes that occur within each polygon. As such, multiple crimes can be recorded at a single spatial location, leading to uncertainties for here presented study.

4.2.3 Data Pre–Processing

Having identified these limitations, I first cleansed the telecommunication data, so to remove inconsistent entries (i.e., footfall count per area different from the sum of footfall counts broken down by gender, type or age). I further pruned grid cells that fell outside the Greater London area. This caused 1.8% of the raw telecommunication data to be removed.

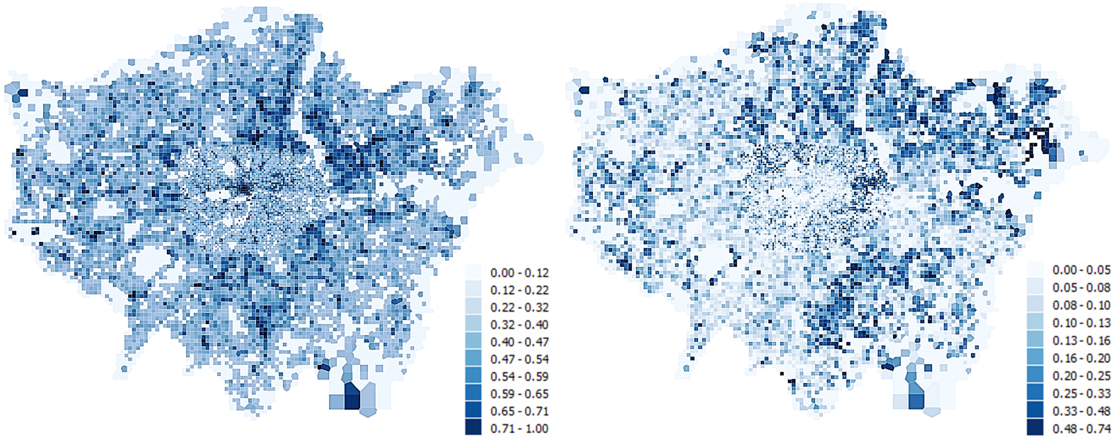
In order to correlate people dynamics and crime data within an urban environment over time, I then needed to define a common spatio–temporal unit of analysis for both datasets. In terms of *spatial* unit of analysis, I operated at the level of grid cells defined by the telecomm operator as this was the coarser granularity. As mentioned before, these are rather fine–grained cells, varying from 210 $\times$ 210 meters for inner London, to 425 $\times$ 425 meters for outer London. As crime data is recorded in terms of latitude/longitude coordinates, the spatial association of crime data to grid cells was straightforward. For each grid cell, I can thus count the total number of crimes that took place there. As advised by the Jill Dando Institute – I visited the institute on several occasions due to the "transport and crime" research group meetings, offering me a platform to present and discuss my work with crime scientists working in the field (Kate Bowers, Reka Solomosi, Matthew Ashby, Aiden Sidebottom Tom Cohen, Sarah Wise) – I break down counts by crime type, distinguishing *street crime*, covering crime most likely happening on the streets (e.g., antisocial behavior, drugs, robbery and violent crime – a total of 47,238 entries), and *home crime*, including crime types happening most likely indoors (e.g., on burglary, criminal damage and arson, other theft and shoplifting – a total of 36,288 entries). In terms of *temporal* unit of analysis, I needed to align telecomm data, captured hourly, with crime data, captured monthly. To do so, I computed average footfall counts per area per month; to reduce variance, I aggregated separately day–time hour slots (8AM-8PM) and night–time hour slots (8PM-8AM), as well as weekdays vs. weekends. For each grid area, I thus ended up with four footfall count averages. Having cleansed the data and defined a common spatial and temporal unit for analysis, I am now able to define the metrics I will use in my quantitative analysis.

4.2.4 Hypotheses & Metrics

Crime count and crime activity. To begin with, I need to quantify crime per spatio–temporal unit of analysis. For each area i, I consider two complementary metrics: crime count $CC(i)$, and crime activity $CA(i)$. The former simply counts the number of crimes that have taken place in area i; since most of the areas under study have comparable size, I may consider $CC(i)$ as a way of measuring crime normalized by area size. Areas have similar sizes, but not similar population density. To investigate possible differences caused by population density, I use $CA(i)$ to quantify crime normalized by population density instead; we can consider this metric as an indicator of the probability of being victim of a crime. I can compute crime activity $CA(i)$ by dividing the number of crimes in an area $CC(i)$ by the estimated population $P(i)$ present in area i. The number of crimes per area $CC(i)$ is available in my pre–processed crime dataset; as for the number of people present in the area, I considered all people present in area i in the 3 weeks covered by my phone call dataset. Since the crime dataset and telecommunication dataset covered different timespans (8 weeks for the former, 3 weeks for the latter), I multiplied by 3/8 so to have the average number of crimes per person in one week:

$$CA(i) = 3/8 \cdot \frac{CC(i)}{P(i)}$$

Figure 4.2 shows the spatial distribution of crime count and crime activity over Greater London (the darker the shade of blue, the higher the $CC(i)$ and $CA(i)$ values). As shown, crime count $CC(i)$ is found to be higher in the centre of London, with some other hotspots spread out all over the city (Figure 4.2a), whereas crime activity $CA(i)$ (that is, crime count normalised by people present in that area) is much higher outside inner London (Figure 4.2b). Having defined a metric that captures crime per spatio-temporal unit of analysis, I next define metrics that act as proxies for urban crime theories linking people dynamics with crime count and crime activity. Selected theories are examples and show the application of the method which can be used to evaluate other theories as well with relevant data at hand. For this study I have a total



(a) crime count **(b)** crime activity

Figure 4.2: Choropleth maps showing crime count CC (left) and crime activity CA (right) all over Greater London for Dec 2012-2013, where the darker the shade of blue, the higher the crime rate in that area

of six metrics and associated hypotheses ($H1$ to $H6$).

H1 - Diversity of people. According to Jacobs, diversity of functions in an area supports the area's safety, as it attracts a greater diversity of people at different times that collectively act as 'eyes on the street'. Jacobs (1961) points out in her examples the importance of age diversity. Newman (1972), on the contrary, suggests that high diversity of people in an area provides opportunities for crime to happen through anonymity. However, the two theories do not describe the term 'diversity' in further detail. From my telecommunication dataset, I am able to extract one metric of diversity, relative to age. For each area under examination, I have a footfall count breakdown relative to age in terms of these age groups: 0-20, 21-30, 31-40, 41-50, 51-60, 60+. I thus computed age diversityD_a as the Shannon–Wiener diversity index (Shannon and Weaver, 1949) over these counts. The Shannon diversity index is a measure that reflects how many different entries there are in a data set and the value is maximized when all entries are equally high. When correlating this metric with crime, according to Jacobs I would expect areas with higher age diversity to be safer than others, while following Newman's theory I would expect the opposite.

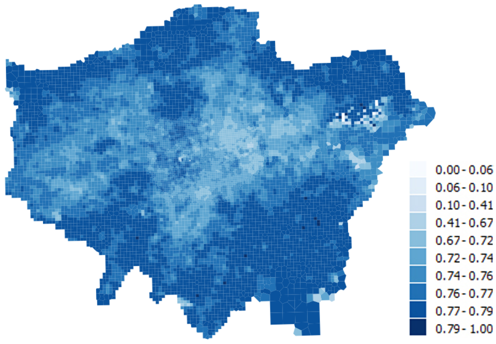
H2 - Ratio of visitors. According to my reviewed theories, there are opposite opinions about the contribution towards crime of a high ratio of visitors to an area. Jacobs points out their importance for 'eyes on the streets', while Newman suggests that a high ratio of visitors actually brings crime to an area as a result of anonymity. To explore these apparently contrasting theories, I quantify the ratio of visitors R_v (relative to total footfall count) per area, and will then correlate these values with crime metrics. Following Jacobs, I would expect to have less crime where there are more visitors, whereas following Newman I would expect the opposite.

H3 - Ratio of residents. A high number of residents in an area is strongly supported by Newman's territorial approach of creating 'defensible space' to reduce crime. Jacobs mentions residents as a less important factor for the 'natural surveillance' theory compared to shopkeepers, as residents provide less attention for street level activities. To validate Newman's theory, I compute the ratio of residents R_r compared to the overall population, and correlate them with crime metrics. According to Newman, I would expect a high ratio of residents in an area to correlate with less crime.

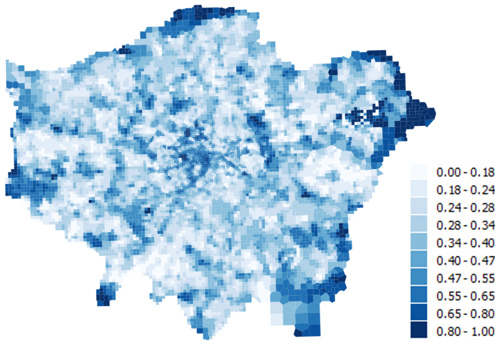
H4 - Ratio of workers. Jacobs suggests that a high variety of functions in an area supports urban safety, pointing out the importance of shops in an area, as shop–keepers and people who work in an area provide 'natural surveillance'. I will test the hypothesis by computing the ratio of workers R_w compared to the area's overall population for each area, and compute correlations with crime metrics. According to Jacobs' theory, I would expect to have less crime in areas with a higher ratio of workers.

H5 - Ratio of female population. Felson and Clarke suggest that a high ratio of women on the street is a positive sign of urban safety, as they act as 'crime detractors'. To test this, I will compute the ratio of female population R_f compared to the overall population for each area, and correlate the values with crime metrics. I would expect a lower crime activity in areas with a higher ratio of females according to the theory.

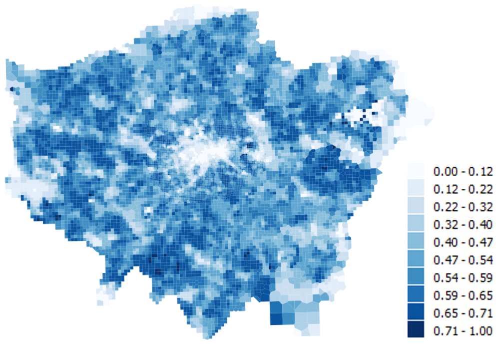
H6 - Ratio of young people. According to Felson and Clarke, a higher ratio of young people leads to more criminal incidents in an area, as they show a higher aggression potential compared to elder people. I defined my young population group as those falling in the 0-20 and 21-30 age groups in my telecommunication dataset, according to (UNDESA, 2014). I then compute the ratio of young (R_y) population relative to the area's overall population, and correlate it with the crime activity. In this case, the hypothesis is that areas with a higher ratio of young people also have higher crime rates.



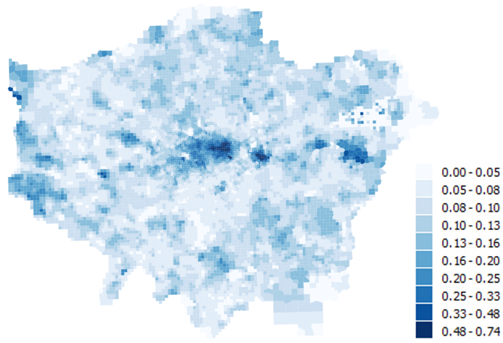
(a) age diversity



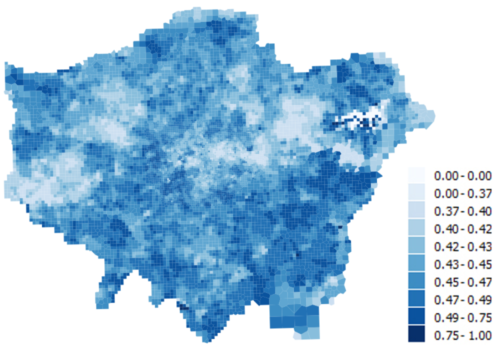
(b) ratio visitors



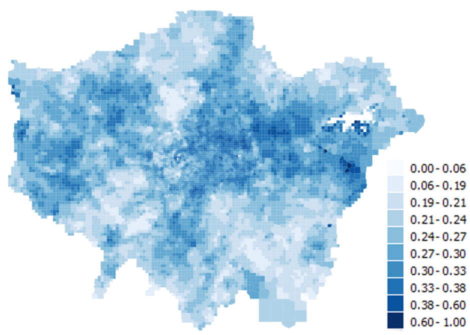
(c) ratio residents



(d) ratio workers



(e) ratio female



(f) ratio young

Figure 4.3: Choropleth maps of my six metrics, where the darker the shade of blue, the higher the value of the metric

Summary of Metrics. Figure 4.3 illustrates the distributions of my six metrics across Greater London as choropleth maps. We observe that the population's age diversity (Figure 4.3(a)) is generally low for Inner London, while it increases towards the edges. A high ratio of visitors is found in the centre of London (Figure 4.3(b)), which offers most points of interest as attractions and retail. Additionally I also found a high ratio of visitors in some parts of the edges towards the north and the east, which looks at first glance surprising. However, this might result from people using their mobile phones while driving along the M25, London's circular motorway, which occupies parts of the areas included in our dataset in these locations. Ratios of residents (Figure 4.3(c)) and workers (Figure 4.3(d)) show a clear opposite picture between them: while workers concentrate in the central business districts, residents are found to be more widespread in less central boroughs. In Figure 4.3(e) we observe generally a higher female population ratio for the south of London, compared to the north. Finally, Figure 4.3(f) shows a higher concentration of young population in the centre of London spreading out towards the east, which is known to be popular among young people.

4.2.5 Analysis

Having defined metrics for crime count, crime activity and the six proxies relating to selected urban crime theories, I can now proceed to evaluate selected crime theories. Focussing on the descriptive, and less on the predictive aspect, I applied correlation analysis to the data. The major challenge of my approach was to manage the spatial autocorrelation present in my datasets. Spatial autocorrelation is rather common when studying spatial processes, whereby observations captured at close geographic proximity appear to be correlated with each other, either positively or negatively, more than observations of the same properties at further distance (Legendre, 1993). This is the direct quantitative demonstration of Tobler's First Law of Geography, which states that '*everything is related to everything else, but near things are more related than distant things*' (Tobler, 1970). Spatial autocorrelation violates the assumption that observations are independent; as such, common correlation analysis techniques that use Pearson, Spearman or Kendall coefficients to explore relationships between variables cannot be applied. To address this issue, I use the Tjostheim correlation index instead (Tjostheim, 1978; Hubert and Golledge, 1982); this index can be seen as an extension to Spearman and Kendall coefficients, to explicitly account for spatial properties in my data. All results presented in the next section are thus to be interpreted as correlations r_t computed between crime count $CC(i)$, crime activity $CA(i)$ and the six metrics $H1 - H6$, using the Tjostheim correlation index.

I then built an Ordinary Least Square (OLS) regression model to explore the relative importance of each of defined metrics to urban crime as well as to assess the overall model's performance to describe the variance in recorded crime. As I found crime val-

Hypothesis	Variable	crime count $CC(i)$ Total Crime	Street Crime	Home Crime	crime activity $CA(i)$ Total Crime	Street Crime	Home Crime
H1: diversity of people	D_a	**-0.27**	**-0.26**	**-0.23**	**-0.12**	**-0.14**	-0.10
H2: ratio of visitors	R_v	**-0.20**	**-0.20**	**-0.17**	**-0.28**	**-0.26**	**-0.23**
H3: ratio of residents	R_r	**0.17**	**0.19**	**0.14**	**0.27**	**0.26**	**0.21**
H4: ratio of workers	R_w	0.09	0.07	0.09	0.02	0.02	0.03
H5: ratio of females	R_f	-0.02	-0.02	-0.01	**0.16**	**0.14**	**0.16**
H6: ratio of young	R_y	**0.31**	**0.31**	**0.25**	**0.13**	**0.17**	**0.10**

Table 4.3: Tjostheim Correlations r between crime metrics (crime count and crime activity) and individual variables; bold results are statistically significant with p-value < 0.01

ues in my data not normally distributed, it was first necessary to transform the data by applying a square–root transformation to the values. Also, I tested the model for multicollinearity by computing the variance inflation factor (*VIF*) to ensure that variables are not affecting each other. Next I will present outcome of these analysis.

4.3 Results

4.3.1 Results for Greater London

Table 4.3 presents the Tjostheim correlation coefficients between my two crime metrics ($CC(i)$ and $CA(i)$) and each variable introduced in the previous section. Bold results are statistically significant with p-value < 0.01, which is providing us with strong evidence against the null hypothesis at a confidence level of 99%. Note that the same correlation signs were found both when using crime count and crime activity, with only relatively small changes in actual correlation values. I interpret this as an indication of the robustness of my proposed metrics. The findings discussed below apply to both crime metrics used.

H1: Diversity of people. I found significant negative correlations between diversity of age and crime, both for total crime ($r_t = -0.27$ for CC and $r_t = -0.12$ for CA) and for street crime ($r_t = -0.26$ for CC and $r_t = -0.14$ for CA); for home crime, I found significant results only for the correlations with CC ($r_t = -0.23$) whereas for CA the p-value was found to be greater than 0.01 so the result is not statistically significant. These findings seem to support Jacob's theory of 'natural surveillance', where she linked different age groups in the same area to a variety of activities taking place in the same space, and this was further associated to less crime.

H2: Ratio of visitors. I found a significant negative correlation between the ratios of visitors (R_v) to an area and crime. For total crime, I found $r_t = -0.20$ for CC and

$r_t = -0.28$ for CA; for street crime, $r_t = -0.20$ and $r_t = -0.26$ respectively; and for home crime $r_t = -0.17$ and $r_t = -0.23$ (second row of Table 4.3). In all three cases, a higher ratio of visitors is linked to lower crime. These findings again support Jacobs' theory of 'eyes on the street', with a consequent increase in the levels of safety of an area where visitors concentrate.

H3: Ratio of residents. Focussing on residents, I found a positive correlation between the ratio of residential population (R_r) in an area and crime. Newman's theory of 'defensible space' suggests that an increased ratio of residents is linked to urban safety, by clearly separating spaces for visitors from spaces for residents. However, my findings do not seem to support this. In fact, results show that a high ratio of residents is statistically correlated with crime (from a minimum of $r_t = 0.14$ for home crime correlated with crime count CC, to a maximum of $r_t = 0.26$ for street crime and crime activity CA (third row of Table 4.3).

H4: Ratio of workers. Contrary to Newman, Jacobs suggests that residents are less involved with natural surveillance compared to, for example, shopkeepers, as they provide less attention to what is taking place around. Jacobs suggests to look at the relationship between the ratio of working people (R_w) in an area and crime instead. In particular, she posits that a high number of functions, especially shops, leads to increased safety as they attract people and support 'natural surveillance'. Unfortunately, my results do not help shed light into this controversy, as they are not statistically significant (fourth row of Table 4.3).

H5: Ratio of female population. A surprising result is found in the positive correlation between the female population (R_f) and crime activity CA in an area ($r_t = 0.16$ for total crime, $r_t = 0.14$ for street crime and $r_t = 0.16$ for home crime – fifth row of Table 4.3), though correlations with crime count CC were found to be not significant . This result shows the opposite of Felson and Clark's theory, suggesting that a higher ratio of female population in London is actually statistically correlated to a higher crime activity in an area. However, I should note a limitation of my metric in this case: in fact, R_f represents the overall ratio of female population for an area (residents, workers, or visiting), and not only the ratio of female population on the streets, so this result could have been affected by a relatively poor metric.

H6: Ratio of younger population. Finally, I have computed the ratio of young people (R_y) per area and correlated it with crime. Findings show a positive correlation between the two (from a minimum of $r_t = 0.10$ for home crime and crime activity CA, to a maximum of $r_t = 0.31$ for total/street crime and crime count CC – last row of Table 4.3). This result would support Felson and Clarke's theory that a higher proportion of young population is associated with more crime in an area.

Having identified correlations between my defined metrics and crime, I next built an Ordinary Least Square (OLS) regression model to explore the relative importance of each of them to crime. To ensure that there are no correlations between the independent variables, increasing the variance of the regression coefficient, I tested first for multi–collinearity by computing the variance inflation factor (*VIF*). The *VIF* indicates how much an estimated regression coefficient's variance increases if predictor variables are correlated. According to the literature (Hair et al., 1995; Neter et al., 1989), a common rule of thumb followed by most practitioners suggests that a *VIF* = 10 or higher indicates that there is severe or serious multi–collinearity between variables, that should be removed from the model. Other literature is stricter, and suggests maximum *VIF* = 5 (Rogerson, 2001) and even *VIF* = 4 (Pan and Jackson, 2008).

In my case, I found severe multi–collinearity for the variables of ratio of visitors R_r (*VIF* = 12.85) and ratio of residents R_r (*VIF* = 18.09) in relation to crime count CC and crime activity CA for all crime groups (total crime, street crime and home crime). The other variables, such as diversity of people D_a (*VIF* = 2.12), ratio of workers R_w (*VIF* = 4.20), ratio of females R_f (*VIF* = 1.81) and ratio of young population R_y (*VIF* = 1.94) were within defined boundaries of acceptance, as suggested by the literature. Based on this finding, I computed a stepwise linear regression to define which variables to keep for next steps. Outcome shows best fit for the model without the variable of ratio of residents R_r (see adjusted–R^2 in Table 4.4).

In Table 4.4 I present outcomes of my models with crime count $CC(i)$ and crime activity $CA(i)$ as dependent variables, and my five defined metrics as independent variables. I show Beta–coefficients β, indicating the relative importance of each variable to crime, and adjusted–R^2, indicating each model's performance in explaining crime variance.

	crime count $CC(i)$						crime activity $CA(i)$					
	Total Crime		Street Crime		Home Crime		Total Crime		Street Crime		Home Crime	
Variables	p-val	β	p-val	β	p-val	β	p-val	β	p-val	β	p-val	β
H1: diversity of people	***	-0.31	***	-0.28	***	-0.24	**	-0.15	***	-0.15	***	-0.13
H2: ratio of visitors	***	-0.25	***	-0.24	***	-0.20	***	-0.26	***	-0.24	***	-0.24
H4: ratio of workers	*	0.10		0.04	*	0.12	*	0.09		0.08	*	0.11
H5: ratio of females		-0.05		-0.03		-0.06		0.05		0.07		0.04
H6: ratio of young	***	0.28	***	0.22	**	0.27	***	0.18	***	0.16	*	0.15
Adjusted–R^2		0.12		0.13		0.10		0.11		0.12		0.08

Table 4.4: Beta–coefficients β between crime metrics (crime count and crime activity) and individual variables, indicating the relative importance of each variable, and adjusted–R^2, indicating the model's performance to predict crime. p<0.001 '***' p<0.01 '**' p<0.05 '*'.

The coefficients show same signs and similar results as my correlation results, when using crime count and crime activity, with only relatively small changes in actual values. We observe that a high diversity of people and ratio of visitors have positive influence on urban safety, as they show relatively high and negative coefficients. For diversity of people, this especially applies to the crime count metric $CC(i)$ ($\beta = -0.31$ for total crime, $\beta = -0.28$ for street crime and $\beta = -0.24$ for home crime), while for crime activity $CA(i)$ results show a slightly lower importance ($\beta = -0.15$ for total crime, $\beta = -0.15$ for street crime and $\beta = -0.13$ for home crime). For ratio of visitors, results show similar importance for both of my crime metrics: crime count $CC(i)$ ($\beta = -0.25$ for total crime, $\beta = -0.24$ for street crime and $\beta = -0.20$ for home crime), crime activity $CA(i)$ ($\beta = -0.26$ for total crime, $\beta = -0.24$ for street crime and $\beta = -0.24$ for home crime).

On the other side, we observe that a high ratio of workers and young population have negative influence on urban safety, as they show positive coefficients. For ratio of young population, this especially applies to the crime count metric $CC(i)$ ($\beta = 0.28$ for total crime, $\beta = 0.22$ for street crime and $\beta = 0.27$ for home crime), while for crime activity $CA(i)$ results show a slightly lower effect ($\beta = 0.18$ for total crime, $\beta = 0.16$ for street crime and $\beta = 0.15$ for home crime). For ratio of workers, results show similar importance for both of my crime metrics: crime count $CC(i)$ ($\beta = 0.10$ for total crime, $\beta = 0.04$ for street crime and $\beta = 0.12$ for home crime), crime activity $CA(i)$ ($\beta = 0.09$ for total crime, $\beta = 0.08$ for street crime and $\beta = 0.11$ for home crime).

In summary, H1, H2 and H6 are the metrics that show highest effects on urban crime (highest β). However, looking at the outcome values for adjusted–R^2 of my

Variable	Min		1st Qu.		Median		3rd Qu.		Max
D_a	-0.51	**	-0.27	***	-0.20	**	-0.12		0.23
R_v	-0.53	**	-0.30	***	-0.20	***	0.00	*	0.18
R_w	-0.28	***	-0.02	**	0.09	*	0.17	*	0.44
R_f	-0.28	*	-0.08	***	0.03	*	0.17	*	0.47
R_y	-0.18		0.18		0.24	***	0.40	**	0.54

Table 4.5: Summary statistics of the Tjostheim correlations between total crime count *CC* and each individual variable on the 32 London boroughs. Stars indicate the percentage of Tjostheim correlations that are statistically significant in each quartile (p-values < 0.01): 0% ‘ ’ 25% ‘*’ 50% ‘**’ 75% ‘***’ 100%

Variable	Min		1st Qu.		Median		3rd Qu.		Max
D_a	-0.41	***	-0.19	***	-0.11		0.01	*	0.45
R_v	-0.57	***	-0.34	**	-0.27	***	-0.18	**	-0.03
R_w	-0.32	***	-0.08		0.02	*	0.11	**	0.39
R_f	-0.18		0.02	*	0.15	***	0.25	**	0.47
R_y	-0.41	*	0.01		0.08	**	0.22	**	0.45

Table 4.6: Summary statistics of the Tjostheim correlations between total crime activity *CA* and each individual variable on the 32 London boroughs. Stars indicate the percentage of Tjostheim correlations that are statistically significant in each quartile (p-values < 0.01): 0% ‘ ’ 25% ‘*’ 50% ‘**’ 75% ‘***’ 100%

models, we observe a relatively low model fit, suggesting we can not explain crime suffiently using my defined metrics only. With my metrics, I am able to explain up to 13% only for all crimes within Greater London (crime count $CC(i)$–model on Street Crime: adjusted–$R^2 = 0.13$). As Greater London is a very large and complex city, I next apply my analyses on smaller areas of London, more specifically on borough level, to see if there exist different theories at play, in different areas.

4.3.2 Zooming in at Borough Level

I have shown how one may use my proposed methodology to quantitatively study the validity of certain urban crime theories at scale. However, one may wonder whether the chosen scale (that is, the whole metropolitan area of London) is appropriate for this type of investigations. As mentioned before, London is a very large and complex city, composed of many different neighbourhoods. Choosing the whole of London as a single context to study urban theories may thus hide the fact that, in practice, different theories and correlations may hold in different London neighbourhoods. Indeed, past studies by Jacobs (1961) and Newman (1972) performed at neighbourhood level, never at a big geographic scale, such as Greater London as a whole.

As my proposed methodology is not prescriptive of a particular size of geographic area, I have repeated the analysis, this time separately considering the 32 administrative boroughs in which London is divided. As the source data is available at grid cell size of up to 400 by 400 metres, it prevented a more fine grained level analysis, such as on ward or LSOA level. I assigned grid cells to borough boundaries according to their centroids. Table 4.5 shows summary statistics of the correlations between crime count CC and each variable previously defined, as they vary across boroughs; Table 4.6 shows results obtained when using crime activity CA instead. By looking at these new results, and by comparing them with those in Table 4.3, I note that all the individual variables that were (positively or negatively) correlated to crime activity in the whole city of London, now show considerably higher (in positive or in negative) correlations in at least half of the 32 London boroughs. This indeed suggests that this smaller unit of analysis can be more appropriate to investigate the validity of urban crime theories. For those metrics for which I did not find significant statistical results when considering the whole of London, I now find significance in certain areas. For instance, my findings reveal that a quarter of London boroughs have a significant negative correlation between the ratio of working population (R_w), and both crime count CC ($-0.28 > r_w > -0.02$) and crime activity CA ($-0.32 > r_w > -0.08$), whereas for Greater London correlations of the same variable were found not to be significant (CA: $r_w = 0.02$, CC: $r_w = 0.09$). Interestingly, the results at borough level also show that, for another quarter of London boroughs, R_w is actually significantly and positively correlated with crime activity CA ($0.11 > r_w > 0.39$) and crime count CC ($0.17 > r_w > 0.44$) instead. These findings suggest that different, possibly conflicting theories may hold in different parts of the same metropolitan city; using my method, it is possible to investigate whether a theory holds at the full city scale or not. If not, the method also helps social science researchers identify the sub-areas that require further qualitative investigation.

4.4 Discussion

4.4.1 Summary

In this chapter, I have presented a method to investigate theories of urban crime and people dynamics in a quantitative way. The method requires access to two sources of information: crime data records and records about people presence in the built environment. From the former, I extracted two metrics of crime, crime count $CC(i)$ and crime activity $CA(i)$. From the latter, I extracted metrics that act as proxies for urban crime theories. Using correlation and regression analysis, I have shown it is now possible to quantitatively investigate urban crime theories at large geographic scale and frequent intervals.

Supported by the ongoing open data movement, an increasing amount of crime data for cities in different parts of the world is freely available and can be used for these purposes. Telecommunication data on the other hand is more difficult to access, but a variety of data mining challenges, such as the Data for Development challenge (Orange, 2013) and the Big Data Challenge (TelecomItalia, 2012) show a clear trend of mobile phone providers towards making their data available to the public. This development suggests that the proposed methodology will become increasingly applicable in the next years.

4.4.2 Limitations

This work suffers from a number of limitations. First, my method is dependent on quality and extent of provided data. The temporal unit of analysis used in the two datasets at hand was different (i.e., crime data was recorded on a monthly basis, while foot–counts were recorded on a hourly basis). This required a data–processing step that forces me to operate at the coarser level of granularity. This inevitably kept interesting questions unanswered. As previous studies suggest, different crime types follow different spatial and temporal patterns (Felson and Poulsen, 2003); if I had access to crime timestamps, I would have been able to explore the relationship between people dynamics and crime in a more fine–grained manner. Also at this point, as this work focuses on the testing of established theories on large scale using correlations, it does not imply any causation between crime types and people dynamics. However, being used in triangulation with additional research coming from social, criminological or urban studies, it offers a new perspective on the subject that can help to uncover these causation relationships.

Furthermore, these findings are based on mobile phone data collected by a single mobile phone provider. Being one of the major mobile phone providers in the UK with almost 25% market share in 2013, my dataset covers a high number and variety of people, but leaves a grey space for people using other providers or PayAsYouGo options that are excluded from the data. For those people covered from my dataset, it stays unclear how the provider categorized them as resident, worker or visitor which could provide a more detailed insight. By including additional datasources, as for instance urban topology or land–use data, the ratio of workers could be discussed in more detail through establishing relationships to urban venues.

Besides the number of people that are being excluded from this data, the procedure of localizing mobile phones brings up limitations in terms of spatial accuracy. Defined by grid–cells rather than geo–locations, our data provides an image that does not allow us to pin–point its users, which is a problem especially in a diverse city such as London where situations can change from one road to the other. It also does not allow us to differentiate between indoor and outdoor users. Furthermore, mobile phones are localized using cell–tower triangulation, using their distances to the cell–towers they are connected to, to compute the approximate location, that might affect our findings.

The open–source crime data set includes flaws in terms of its spatial accuracy as it provides geo–locations of map points pre–defined by London police (PoliceUK, 2013b), but not about the location where the crime actually has happened.

Note that these limitations pertain the datasets used, and not the method proposed. Here we have shown how to apply the method, as if the data is flawless. While actual results on the validity of the reviewed urban crime theories for the case study of Greater London would have to be revisited should more accurate and complete datasets become available. At this point my defined proxies are simplifications of actual theories. For instance, in her work, Jacobs (1961) suggests the importance of diversity – especially diversity of age – in relation to crime. However, diversity can be based on other variables as well, such as people's ethnicity or social status. With my dataset at hand, I was able to discuss the variable of age diversity only. It stays unclear however, how results would change if I use different metrics to capture, as in this case, diversity.

4.4.3 Implications

The method I have proposed has both practical and theoretical implications. From a practical standpoint, tools can be built on top of it, for the benefit of different stakeholders, as citizens, administrators and city planners. To illustrate what such a tool would look like, I built an Ordinary Least Square (OLS) regression model for each of the 32 boroughs in Greater London separately, as well as for the whole of London. For each regression model, I analysed the adjusted–R^2 value, to understand the extent to which the built model was capable of 'explaining' crime variance. I found that, for a model that considers Greater London as a whole, the adjusted–R^2 value is 0.12. However, when I built such model per borough, I was capable of reaching an adjusted–R^2 between 0.21 and 0.32 for a quarter of the boroughs. These results offer opportunities for instance to city administrators and police to improve crime prediction models by combining my method with other sources describing the urban environment, such as the built (Hillier and Hanson, 1984) and the social environment (GLA, 2011).

From a theoretical standpoint, the method offers social science researchers a new way to investigate past crime theories, as well as develop new ones. I have shown how to use the method to explore past theories for the city of London. The same method could be used for a multitude of cities around the world, so to advance knowledge in terms of the contexts within which past theories hold. The method can also be re–applied over time, on newly available data streams, to detect possible changes that call for social scientists to refine past theories or develop new ones. Even when looking at the single city of London in a single period, I have shown that some theories do not hold across all boroughs, thus calling for deeper qualitative investigations in selected areas. I foresee the proposed quantitative method to be used in conjunction with qualitative methods, during alternate phases of theory development and evaluation.

4.4.4 Conclusion

In this chapter I have presented a method enabling researchers to quantitatively test well established urban crime theories based on qualitative studies. As defined in Chapter 2, besides crime, fear of crime has become an increasing problem for the broad population (Brown and Polk, 1996; Oc and Tiesdell, 1997). Describing a perception, such as a lack of feeling of safety by the urban population in the environment, fear of crime stands in close relation to crime activity (Doran and Burgess, 2012) but furthermore is not restricted to space and time (Matei et al., 2001). As these circumstances lead to serious problems for a city's inhabitants and its government, it is important to include fear of crime into the discussion. In the next chapter I will propose a crowdsourcing method to gather safety perceptions towards people on large scale.

Chapter 5

A Crowdsourcing Approach to Evaluate and Develop Fear of Crime Theories towards People

Part of the work presented in this chapter also appeared as a full paper accepted to the 7th International Conference on Social Informatics (SocInfo) 2015) [Acceptance Rate: 23%]:

Traunmueller, M., Marshall, P. and Capra, L., Crowdsourcing Safety Perceptions of People: Opportunities and Limitations, *In Proceedings of the 7th International Conference on Social Informatics (SocInfo)*, December 2015.

5.1 Introduction

In the last chapter I have discussed the relationship between demographic properties of people, such as age, gender, their relationship to a geographic location and crime in an urban environment. However, in Chapter 2 I have identified that, besides crime, fear of crime has become an increasing problem affecting the broad population, especially in cities (Brown and Polk, 1996; Oc and Tiesdell, 1997). In contrast to actual crime activity, fear of crime describes a perception, such as a lack of feeling of safety by the urban population in the environment. Thereby it stands in close relation to crime activity, as there is a higher fear of crime among the population of areas with high crime rates (Doran and Burgess, 2012). Other researchers suggest furthermore that, supported by newsbroadcasting through modern media, such as television and online social networks (Matei et al., 2001), it is not limited only to victims of actual crime, but even more affects the broad population. In doing so, fear of crime is not restricted to space and time, having the potential to be more widespread than crime. Therefore it is seen as an even bigger problem than crime itself (Brown and Polk, 1996; Oc and Tiesdell, 1997): on one hand, fear of crime limits a city's population in their daily actions by avoiding public space that is perceived to be unsafe (Nasar and Fisher, 1993; Warr and Ellison, 2000), leading to reduced quality of urban life (Pacione, 2003). On the other, this results in a problem for the city's sustainability, as avoiding public space leads to a decrease in the city's walkability and increased motor traffic (Warr and Ellison, 2000).

Describing a perception, such as a lack in feeling of safety in an environment, rather than the actual victimization that has been recorded, I can not rely on existing datasets, but have to create one myself first. To collect a large amount of data about safety perception in a short time, I propose an online crowdsourcing approach using images.

The use of images in research. The use of images has been common practise in social science, where they have been used as part of qualitative studies to explore the relationship between people's looks and a variety of factors. Gold (1991), for instance, uses images of Vietnamese refugees sub–populations in the U.S. – ethnic Vietnamese and Chinese–Vietnamese – to study how refugees can determine the others ethnicity and traits they associate with. In doing so, they also studied if there are differences in understanding between the younger, more Americanized, and the older, more traditional generation. To answer these questions, researchers selected a number of images from a large pool that has been used in prior studies, showing people in their every–day environment, as for instance their small local business. The images selection was presented to participants and semi–structured interviews conducted. Results reveal that overall the two refugee groups are able to visually determine the ethnicity of others and that their ethnic boundaries are less important, the longer participants lived in the U.S.

Also discussing differences in visual perception of ethnic subgroups, Uhlmann et al. (2002) uses images of Latin–American people to examine the effect of skin colour on American hispanics and Chileneans towards subgroups within the community, devided in "Blancos" and "Morenos". In doing so, reseachers selected 30 yearbook portraits of students (10 Caucasian, 10 "Blancos" and 10 "Morenos") and used them in an Implicit Association Test (IAT) (Greenwald et al., 1998) (a task that uses the speed of responses to questions as indication for strength of the attitude, such as positive or negative, participants associate to properties, such as ethnic differences). In addition, they also gathered information about the participants as well, as for instance to which of the two groups ("Blanco" or "Moreno) they consider themselves to belong to. Images were shown in the laboratory on computer screens to participants. Results reveal that both American hispanics and Chileneans preferred strongly images showing the lighter–toned types, no matter if they were "Blancos" or "Morenos". They further investigated perception differences towards Hispanics and Caucasians and found severe differences depending on nationality: While Chilenians expressed favour of Caucasians over Hispanics, American hispanics did not favour any of the two groups.

Dasgupta et al. (2000) use photographic images to discuss perception differences towards Caucasian and Black Americans, controlling for familiarity. Findings revealed that images showing Caucasian people were more strongly correlated with positive attitudes, while images showing Black people negative attitudes, no matter how familiar people were.

The same researchers explored in a follow–up study (Dasgupta and Greenwald, 2001) if admiration and fame of people can reduce automated preference towards Black or Caucasian people, as elder and younger people. They presented pictures of either admired or disliked Black or Caucasian people, as gathered from the internet, to groups of participants, who completed afterwards an IAT to detect racial attitudes. Results show that the exhibition of admired Black people had a significant effect and weakend pro–Caucasian attitudes within 24 hours after the study. However, on the long term it did not affect racial attitudes significanly.

Focussing less on ethnic background, but on physicality and proportions of people‘s faces, more recent work (Vernon et al., 2014) tries to detect features automatically on a set of images, to characterize the relationship between appearance and social traits. In their study, reserachers used a pool of 1000 photographic portaits of Caucasian people, that have been rated in a previous study in terms of social traits, and categorized them using factor analysis in terms of 'approachability', 'youthful–attractiveness' and 'dominance'. An artificial neural network was then used to predict social traits according to the facial expression.

The above mentioned studies show the wide spread use of images, both in qualitative and quantitative studies to analyse the impact of people‘s physical appearance and demographic properties, based on people‘s attitude towards them.

As previously identified in Chapter 2, theories from social and criminological sciences have linked differences in perception of safety to demographic properties of the population. Such properties include for instance people’s age (Zako, 2009), gender (Felson and Clarke, 1998) and ethnicity (Day, 1999; Pain, 2001), which matter in both ways: on *who you are*, and *who you see*. However, based on qualitative methods, these works offer detailed insights but are difficult to replicate at scale, across different cultures, and over time due to the cost associated with them. Recent work in computer science suggests online crowdsourcing as a complementary method to gather perceptions of happiness and safety, amongst others, in relation to the built environment (Salesses et al., 2013; Quercia et al., 2014).

With this background, I propose in this chapter a method to study fear of crime quantitatively and at scale, with the same goal as in Chapter 4, to provide methods that support the advancement of theoretical understanding of the subject. I investigate to what extent online crowdsourcing using images can be used by social scientists to validate theories of safety about *people* instead. While works such as Salesses et al. (2013) and Quercia et al. (2014) might not have had the need to differentiate who provided opinions to validate theories about the built environment, I discuss theories where it matters *who* gives the opinions, by collecting demographic information about my respondents in addition to their safety perception scores.

The remainder of this chapter is structured as follows: first I outline the development and deployment of *Streetsmart*, an online crowdsourcing platform developed to gather safety perceptions about people, define hypotheses and describe my analysis steps. I then present the results of my study and discuss its limitations and implications.

5.2 Method

In this section, I describe the method I propose to quantitatively evaluate fear of crime towards people. I start with identifying potential limitations of this approach, based on crowdsourcing as method. I then discuss the image selection and study preparation process and outline our analysis steps.

5.2.1 The challenges of using actively collected data for studying fear of crime towards people

In this study, I borrow a method from social sciences and use it quantitatively on a large scale by presenting images of people via an online platform, specifically to crowdsource safety perceptions towards them. In doing so, there are a number of risks and challenges that will affect and limit my findings and that I need to discuss upfront.

First of all, it is well known that crowdsourcing suffers from demographic user bias (Quattrone et al., 2015). In this study, I collect user demographic variables, such as age and gender, as defined by the literature (Zako, 2009; Felson and Clarke, 1998; Day, 1999; Pain, 2001). In practise I was able to reach sufficient data only for the demographic group of Caucasian and middle–aged participants. Other demographics I was not able to cover using this method. Therefore, the safety perceptions result I was able to collect are limited to the specific group of Caucasian middle–aged participants only. My findings can not be generalized for other demographic groups.

Furthermore, selecting images that are being used to research safety perceptions by the researcher includes flaws related to reliability and representativeness in terms of bias by the researcher‘s background. In this study, I aim to minimize this limitation by including a broader sample for each person type, but I still need to be aware of such bias as image selection was not cross–validated by any researcher of different background.

While being aware of these drawbacks and limitations, the contribution of this work is in the examination of the usage of this method to study safety perceptions towards people on large scale, by discussing properties as defined by qualitative prior work. With this study I show how to apply a common practise from social sciences – the use of images to capture perceptions towards people – quantitatively and on large scale.

Ethic approval. To ensure ethical correctness of the suggested methodology, I sought approval from UCL's research ethics commitee (UCL, 2016) prior to the study. The procedure includes the submission of a detailled description about the study design, the data that will be collected and the purpose of the study. In my case, as I worked with images used on an online crowdsourcing platform, I additionally submitted screenshot images for better understanding. Submitted materials were being discussed by the commitee and changes suggested: in my case, the inclusion of an 'Exit' button on the webpage interface (enabling participants to quit the survey at anytime) and a further blurring of faces shown on the images, using image processing software, to ensure anonymity. After resubmission I received the ethic approval for the study.

5.2.2 Selecting features

First, I decided what features of people the study should explore. To validate existing theories on fear of crime, I can ground my selection in the literature, as defined in Chapter 2, where I defined that age, gender and ethnicity matter to safety perception, both on *who we are* (e.g., a man and a woman may perceive the same person differently), and *who we see* (e.g., the same person may perceive a young or elderly person differently).

However, many other factors about who we see may impact safety perceptions, such as clothing, posture, facial expression, etc. To explore what other factors we could include and hence, what new theories we could define, I conducted an exploratory pilot study first.

Method. I interviewed 21 people living in London, between the age of 21-64 years, including 13 female and 8 male. Their ethnicities included Caucasian, Black, Indian, Asian and Arab, covering all main ethnical groups in London (GLA, 2011). Images of people were shown to participants, who were then asked to talk aloud (Ericsson and Simon, 1980) about their feeling of safety and reasons for their decision. Each session lasted for 30 minutes, answers were audio recorded and transcribed for analysis.

The images used in the pilot study were selected from online repositories under a creative commons license. As well as including different age groups, gender and ethnicities, I included people's appearance in terms of signifiers of religious (Saroglou, 2014), sexual (Clarke et al., 2012) or sub–cultural (Adams, 2008) orientation, the person's facing–direction, differences in fashion, posture and gestures, differences in activity (walking / standing), the number of people in the picture (single / group), mix of people within a group and if the person was concealed or not, as for instance wearing a hood. 138 images of people were subtracted from the image background and placed on a white canvas showing only perspective lines indicating a street, as shown in Figure 5.1. This step was necessary to focus participants on the person only without

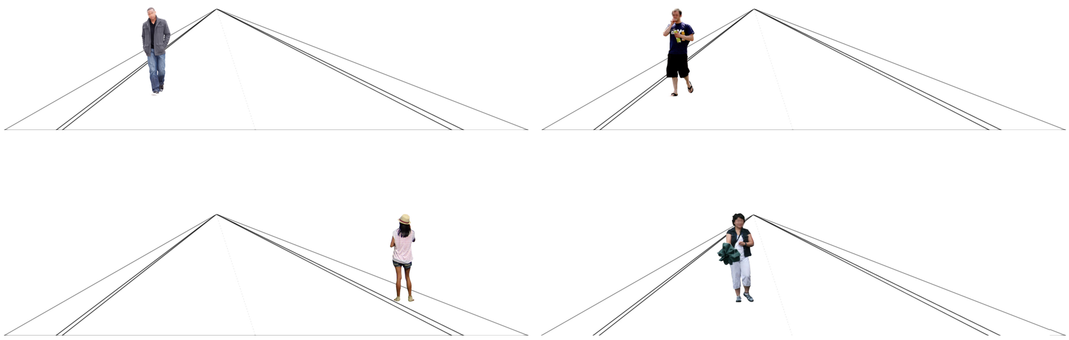


Figure 5.1: Image examples, as used in the pilot study, showing the selected people on white background with black perspective lines, indicating a street.

any distraction from the backdrop, which is known to have an impact on safety perception (Gehl, 2010). People shown in the picture were scaled to appear at the same distance of 6-7 meters to the participant, as in real situations this distance enables us to see enough visual detail in other people to assess whether to be fearful or not (Hall, 1966). Each participant was shown all 138 images and was asked to elaborate about their safety perception towards each person and the reasons for it.

Results. After transcribing audio recordings, I used thematic analysis (Braun and Clarke, 2006) to process the output. I generated initial codes for potential themes represented in the majority of answers and detected patterns emerging over the data. Then I iterated the process and finally was able to describe two main themes to be included in my study beside the already known variables of *age, gender* and *ethnicity*. These were: people's *facing–direction* and if the person's face is *concealed or not*, for example by a hood.

The results from the pilot study show that most study participants felt unsafe when not knowing what the person was up to, which was then linked to trouble. This resulted from concealment of the person's face or if the person faced away from the participants, covering his / her actions. On the oposite, a person facing towards participants, triggered the emotion of unsafety through the feeling of interaction, as participants mentioned. Other properties included in the pilot study, as for instance describing people's sub–cultural or sexual orientation, were found to have less impact on safety perception and hence, were excluded for further steps.

Variables	*Values*	*Type*
Age	Teenager (14 – 28 years) Grown-up (29 – 55 years) Elder (56+ years)	main
Gender	Male Female	main
Ethnicity	Caucasian Black Asian Arab	main
Facing–Direction	towards me away from me not aware of me	sub
Concealed	concealed / not concealed	special

Table 5.1: Table showing breakdown of selected variables.

5.2.3 Selecting images

According to the literature and my findings from the pilot study, I defined 5 variables to take into account when selecting the images: *age, gender, ethnicity, facing–direction* and *concealed*, as shown in Table 5.1. The method is not limited to these 5 variables but could focus on other variables as well that are not covered by this case study. Variables were defined as 'main variables' (as found in the literature, including *age, gender* and *ethnicity*) and as 'sub variable' (*facing–direction*), found in the pilot. Images of concealed people did not allow us to categorize them on the remaining variables, as these properties were hidden, and hence were treated seperately. I broke the main and sub variables down by the values presented in Table 5.1. In total, we obtained 72 different categories including $3\times$ age (0-20, 21-40, 41+), $2\times$ gender (male, female), $4\times$ ethnicity (Caucasian, Black, Asian, Arab) and $3\times$ facing–direction (towards me, away from me, not aware of me).

I then selected pictures of people covering this range from which I subtracted backgrounds, using the online marketplace *Fiverr* (Fiverr, 2014). *Fiverr* is an online platform that enables people to offer their skills for different tasks for $5 USD. The user sets up an account to give instructions and uploads image files. Simple tasks, such as background substraction are suitable for such an approach, leaving not much space for misunderstanding and unexpected results. The returned and processed images of selected people without the background were placed on the neutral white canvas only showing perspective lines in the same distance as mentioned in the pilot study. Each image received a unique ID number and was collected as dataset Di with its

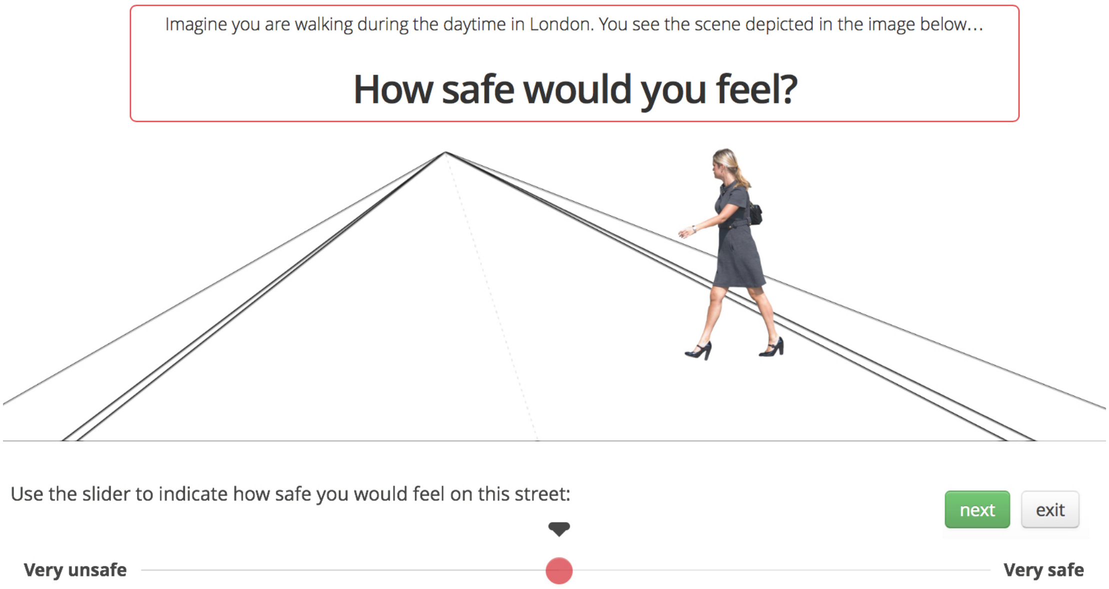


Figure 5.2: Web–based user–interface of *Streetsmart*, showing one image of a person on white background at a time. The user is asked to rate his/her perception of safety according to the image on the slider. By pressing the 'Next' button a new image is presented; by pressing 'Exit' the study can be ended at any time.

variables. To have them look like "normal" people on the street, I avoided images taken by professionals (from magazines for instance) or people posing for the camera and focused on natural snapshots. All pictures used came with a creative commons license. From every category I chose three different images, resulting in a total of 216 images. In this work I focus on individuals only; groups of people are discussed later as future work.

5.2.4 Data collection via online crowdsourcing

The prepared images were used to crowdsource visual safety perception data about people at large scale. To gather large amount of data it was necessary to present them to a broad audience. To do so, I built the *Streetsmart* website, an online crowdsourcing platform enabling people to rate their safety perception according to the image they see. The entry page to the website asked crowdworkers to provide basic background information on who they are. Following the literature, I asked their *gender, age* and *ethnicity*. Furthermore, to be able to understand possible relationships with their geographical / cultural background, I asked which *continent* they come from and if they are a London resident. Then the safety perception survey followed (see Figure 5.2). In accordance to the literature (Salesses et al., 2013; Quercia et al., 2014), I chose a simple and clearly understandable web–interface design, showing from top to bottom the main question, the image and a slider to the participant. A new image was shown by pressing the ‘Next’ button; to quit the survey participants pressed ‘Exit’. I asked crowdworkers to complete 25 safety evaluations using the slider, built as a 100 point Likert–scale. I randomized what pictures to use from a pool of 216, 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.

I obtained ethical approval to conduct the study in September 2014. I then deployed *Streetsmart* from January 2015 until April 2015. I advertised on various social media platforms, such as Facebook (advertisment, Facebook page and group postings), LinkedIn, Twitter and Reddit. Users could share the webpage link via embedded Facebook–Like and Tweet buttons. Throughout these three and a half months, I was able to collect 13,560 votes from 716 users.

A collected time–stamp for each click, indicating how long the user thought about the vote before submitting, gave us a good indication about seriousness of user feedback and hence enabled us to detect meaningless data I could exclude from the study (e.g., because of votes given too quickly to make a serious judgement, as indicated by the time–stamp (Willis and Todorov, 2006)). As it is necessary to know the background information about my participants, I also removed votes from users who did not want to reveal their background, resulting in 8,292 votes from 537 users. The ratio between male (52%) and female (48%) participants was almost even, with most participants aged between 21-40 years (68%), followed by 41+ year olds (25%) and 0-20 year olds (7%). For ethnicity I found a high majority of Caucasian (79%), followed by Asian (7%) and mixed/multiple ethnic groups (5%). 1/3 of my participants were London residents (31%). Overall most participants came from Europe (79%), followed by North America (9%), South America (4%), and Asia (4%).

In Figure 5.3 I show these participant–ratios graphically in pie–charts. While reaching a large number of people in a relatively short time, we can see that some demographic groups were difficult to reach, as for instance Asian people or the group of 0-20 year olds. Looking at this low percentage, we can expect limitations in terms of reaching statistical power for these minorities, which will be included in the discussion. However, looking at London census data (GLA, 2011), we notice that the ethnic

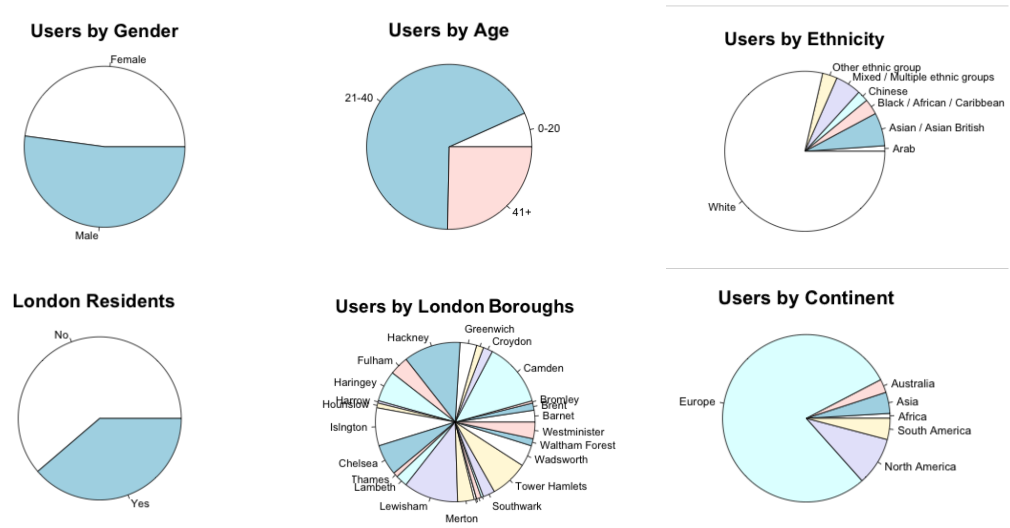


Figure 5.3: Piecharts showing participant ratios for demographical and geographical properties: by gender, by age, by ethnicity, by London residents, by London boroughs, by continent (top to bottom, left to right).

distribution of my data for London residents (29%) is representative of the actual ethnic distribution of the population of London.

5.2.5 Hypotheses

Based on the literature and my pilot study, I formulated the following four hypotheses:

- images showing men would have lower safety ratings then those showing women (Gi);
- ethnicities other than Caucasian and Asian in the image would have lower safety ratings (Ei);
- people in the image looking at the participant would have lower safety ratings than those who were looking away or were not aware of him/her (Fi);
- concealed people would have lower safety ratings than those who were not concealed (Ci).

5.2.6 Analysis

Using the crowdsourced data, I was then able to quantify relationships between participant's safety perceptions and people in the image. To test if our defined demographic properties have different effects on safety perception, I used a factorial repeated Analysis of Variance (ANOVA) with planned contrasts. Effects that were not covered by the hypotheses were explored with post hoc analysis, using the Tukey correction for multiple comparisons (Tukey, 1949). To do so, I first aggregated my collected data to combine the participant information (*gender, age, ethnicity, from London / not from London, where in London, from which continent*), with the properties of the person in the image (*gender, age, ethnicity, facing–direction, concealed / not concealed*) and the safety ratings. As the variable *concealed* does not allow us to define other variables such as the *age, gender* and *ethnicity* of the person in the picture as they stay in the dark, it prevented a factorial comparison so was tested in a separate ANOVA.

Looking at the frequency distribution of ratings, I observed that the data is highly skewed showing a long tail distribution: most people feel very safe in relation to other people, whereas only some feel unsafe. This is a common situation in multi–factorial experiments in human–computer interaction (HCI) that work with Likert responses. In this case, a common transformation, such as the log–transformation, does not work. To transform my data so that I could build an ANOVA model, I used the *Aligned Rank Transform* (Wobbrock et al., 2011) for non–parametric factorial data analysis, which aligns the data in a pre–processing step before applying averaged ranks. With transformed data, I was then able to conduct my analysis.

5.3 Results

5.3.1 Who we see

I started validating excisting theories by analysing main effects. I took the safety score as the dependent variable, and all properties of the user and the person in the image as independent variables. In Table 5.2 I present the mean safety scores M for each variable of the model ($M \in [0, 100]$ with 0 indicating not safe and 100 indicating very safe). Next I summarize the findings.

Gi – *Gender of person in the image*: according to Warr (Warr, 1984) men are perceived as less safe than women. I can support this hypothesis as there was a main effect of gender ($F(1, 7885) = 128.38, p < 0.001$) with men being perceived as less safe than women.

Ei – *Ethnicity of the person in the image*: according to Matei et al. (2001) Caucasian and Asian people are perceived as more safe than other ethnicities. I can support this hypothesis with my findings. I found a significant main effect of ethnicity of the

Hypothesis	*var name*	p–Value	*contained Properties*	*Mean M*
Gender image	Gi	***	male	75.5
			female	79.57
Ethnicity image	Ei	**	Black	76.95
			Asian	79.5,
			Caucasian	78.18
			Arab	75.57
Facing image	Fi	***	away from me	77.08
			towards me	77.09
			not aware of me	78.46
Concealed image	Ci	***	concealed	76.13
			not concealed	77.54

Table 5.2: Table showing breakdown of my discussed hypotheses, with each contained variable and their means of voting scores. Voting ranged from 0 (indicating not safe) to 100 (indicating very safe), p<0.001 '***' p<0.01 '**' p<0.05 '*'.

person in the image ($F(3, 7885) = 21.99, p < 0.01$). Contrasts revealed that situations portraying Black and Arab people were rated less safe than Asian and Caucasian people ($p < 0.001$). Post hoc analysis showed significant differences between situations portraying Asian and Black people ($p < 0.001$) (indicating that Asian people are perceived safer than Black people), Arab and Asian people ($p < 0.01$) (indicating that Asian people are perceived safer than Arab people), Arab and Caucasian people ($p < 0.01$) (indicating that Caucasian people are perceived safer than Arab people) and Black and Asian people ($p < 0.001$) (indicating that Black people are perceived as less safe compared to Asian ethnicities).

Fi – *Facing direction of the person in the image*: from my pilot study, I know that people looking at the participant were perceived as less safe as people looking away or are not aware. There was significant main effect of the facing–direction of the person in the image on safety rating by the users ($F(2, 7885) = 6.50, p < 0.001$). The contrast revealed that people facing towards the participant in the image were rated less safe compared to people facing away or who are not aware of the user ($p < 0.001$). Post hoc analysis showed significant differences in ratings of people who are not aware of ($p < 0.01$), or who are facing away from the participant ($p < 0.05$). This indicates that people turning their back completely to the user are perceived as less safe than those not aware.

Ci – *Concealment of the person in the image*: I found in my pilot study that concealed people were perceived as less safe than those who are not concealed. The outcome of my study ($F(1, 8290) = 156.92, p < 0.001$) supports this hypothesis; concealed people in the image received lower rates than not concealed ones.

Interaction	*var name*	contained properties	*Mean M*	
Age : Ethnicity	$Ai : Ei$	young Asian	80.2	*
		young Black	76.1	*
		young Caucasian	77.3	**
		young Arab	76.8	*
Facing : Ethnicity	$Fi : Ei$	Asian, away	79.1	**
		Black, away	75.9	**
		Caucasian, towards user	77.4	**
		Arab, towards user	74.6	*
Gender : Age	$Gi : Ai$	young male	75.5	***
		grown–up male	74.5	**
		elder male	76.5	***
		young female	79.7	**
		grown–up female	79.7	***
		elder female	79.3	**
Age : Facing	$Ai : Fi$	young, towards user	77.3	**
		young, not–aware	78.8	**
		young, away	76.7	*
		grown–up, towards user	76.1	**
		grown–up, not–aware	78.1	*
		grown–up, away	77.2	*
		elder, towards user	77.9	*
		elder, not–aware	78.5	**
		elder, away	77.3	*
Gender : Facing	$Gi : Fi$	male, towards user	73.7	***
		male, not–aware	77.0	*
		male, away	74.8	**
		female, towards user	79.4	*
		female, not–aware	80.0	*
		female, away	78.3	*
Gender : Ethnicity : Facing	$Gi : Ei : Fi$	Caucasian male, away	74.6	**
		Black male, away	73.3	*
		Asian male, away	77.3	**
		Caucasian female, away	81.4	*
		Asian female, away	81.0	***
		Arab female, away	76.5	*
		Caucasian male, not–aware	79.1	**
		Black male, not–aware	76.3	*
		Asian male, not–aware	77.9	*
		Arab male, not–aware	74.7	*
		Caucasian female, not–aware	79.5	**
		Black female, not–aware	80.3	*
		Asian female, not–aware	81.4	**
		Arab female, not–aware	78.6	*
		Caucasian male, towards user	74.6	***
		Black male, towards user	74.5	***
		Caucasian female, towards user	80.2	**
		Black female, towards user	78.9	*

Table 5.3: Table showing breakdown of my statistically significant interactions with means of safety votes for each variable. Voting ranged from 0 (indicating not safe) to 100 (indicating very safe), $p<0.001$ '***' $p<0.01$ '**' $p<0.05$ '*'.

Interactions. My results for the main effects show to what extent online crowdsourcing can be used to validate theories on urban safety perceptions. However, as in reality such factors are not isolated but happen in interaction with each other, I explored interactions too. In Table 5.3 I present six significant interactions that show clear differences within each group when observed in more detail.

$Ai : Ei$ – There was a significant interaction of *age* with *ethnicity* of the person in the image ($F(6, 7885) = 2.31, p < 0.01$). Simple effects analysis revealed while teenagers were found to be perceived as least safe for situations portraying Black ($p < 0.05$) and Caucasian ethnic groups ($p < 0.01$), they were perceived as most safe situations portraying the groups of Asian ($p < 0.05$) and Arab population ($p < 0.05$) (see Figure 5.4(a)).

While safety perception towards elder (red) and grown–ups (green) follow a similar relationship, teenagers (blue) – generally perceived as least safe – show higher scores for the ethnic groups of Asian and Arab people.

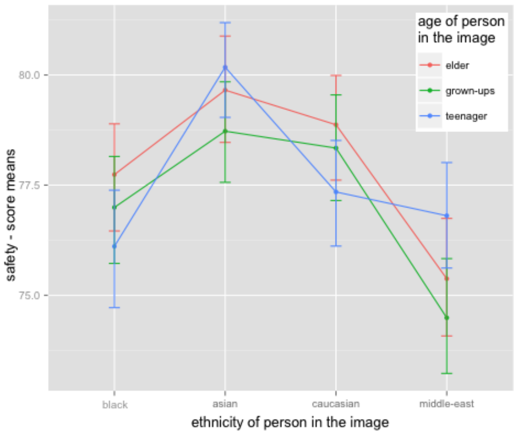
$Fi : Ei$ – There was also a significant interaction between *facing–direction* and the *ethnicity* of the person in the image ($F(6, 7885) = 1.84, p < 0.01$). While situations portraying Caucasian ($p < 0.01$) and Arab people ($p < 0.05$) were perceived as least safe when they looked towards the user, situations portraying Black ($p < 0.01$) and Asian people ($p < 0.01$) were perceived as least safe when facing away from the user (see Figure 5.4(b)).

We observe that generally people who are not aware of the user (green) have been rated as safest. People looking towards the user (blue) – generally rated as least safe – show a higher score, compared to people looking away (red) when they are Black or Asian only.

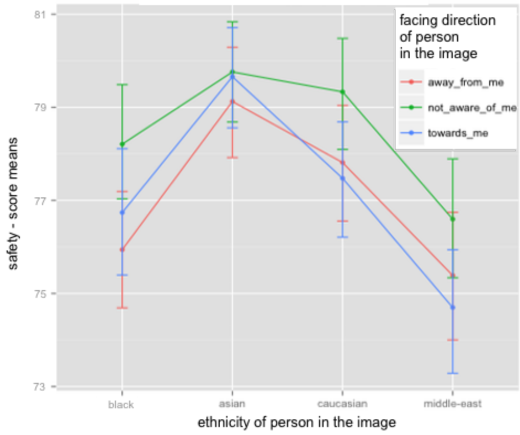
$Ai : Gi$ – I found also a significant interaction between *age* and the *gender* of the person in the image ($F(2, 7885) = 3.77, p < 0.01$). While all age groups of men are perceived as less safe compared to all age groups of women, results show that, within the group of men only, grown–ups ($p < 0.01$) were perceived as less safe than elder men ($p < 0.001$) (see Figure 5.4(c)).

We observe that generally all age groups of women received higher safety scores compared to men. Overall age matters more in the case of men compared to women, where scores were more similar over all ages. Elder people (red) are perceived as most safe and grown–ups (green) as least safe in the case of men.

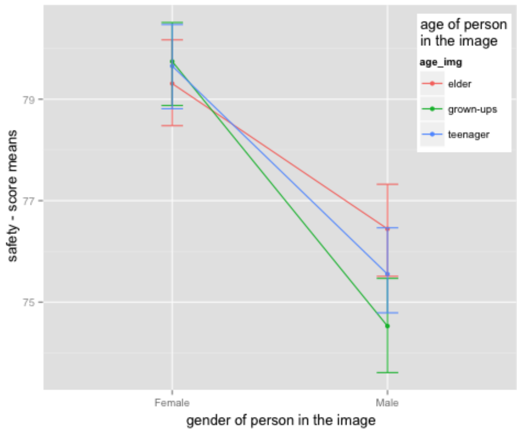
$Ai : Fi$ – There was also a significant interaction between the *age* and *facing–direction* of the person in the image ($F(4, 7885) = 1.84, p < 0.01$). Overall, people were perceived as safest when not being aware, such as elder people ($p < 0.05$), grown–ups ($p < 0.01$) and teenagers ($p < 0.01$). However, different age groups of people were found to be perceived as least safe depending on different facing–directions: While



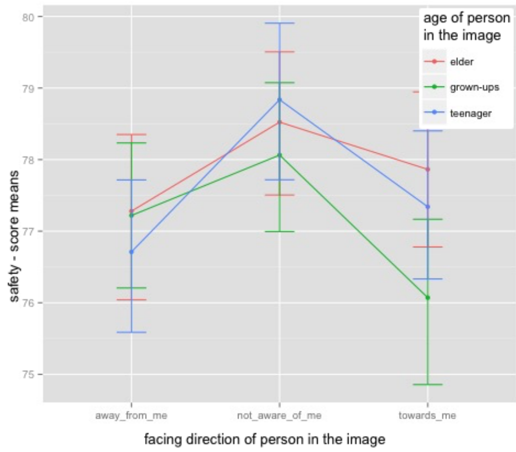
(a) age : ethnicity



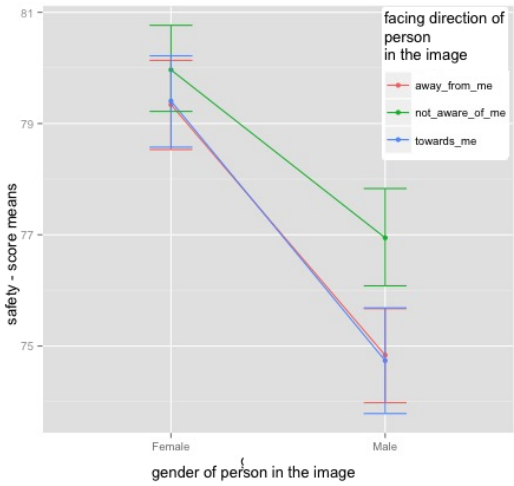
(b) facing–direction : ethnicity



(c) age : gender



(d) age : facing–direction



(e) gender : facing–direction

Figure 5.4: Interaction plots (from top left to bottom): (a) age and ethnicity, (b) facing–direction and ethnicity, (c) age and gender, (d) age and facing–direction, (e) gender and facing–direction of the person in the image. y–axis shows means of perceived safety scores, x–axis shows the variable 1 and colours indicate variable 2 of each interaction.

teenagers were perceived as least safe when they looked away from the user ($p < 0.05$), grown–ups were perceived as least safe when looking towards the user ($p < 0.01$) (see Figure 5.4(d)).

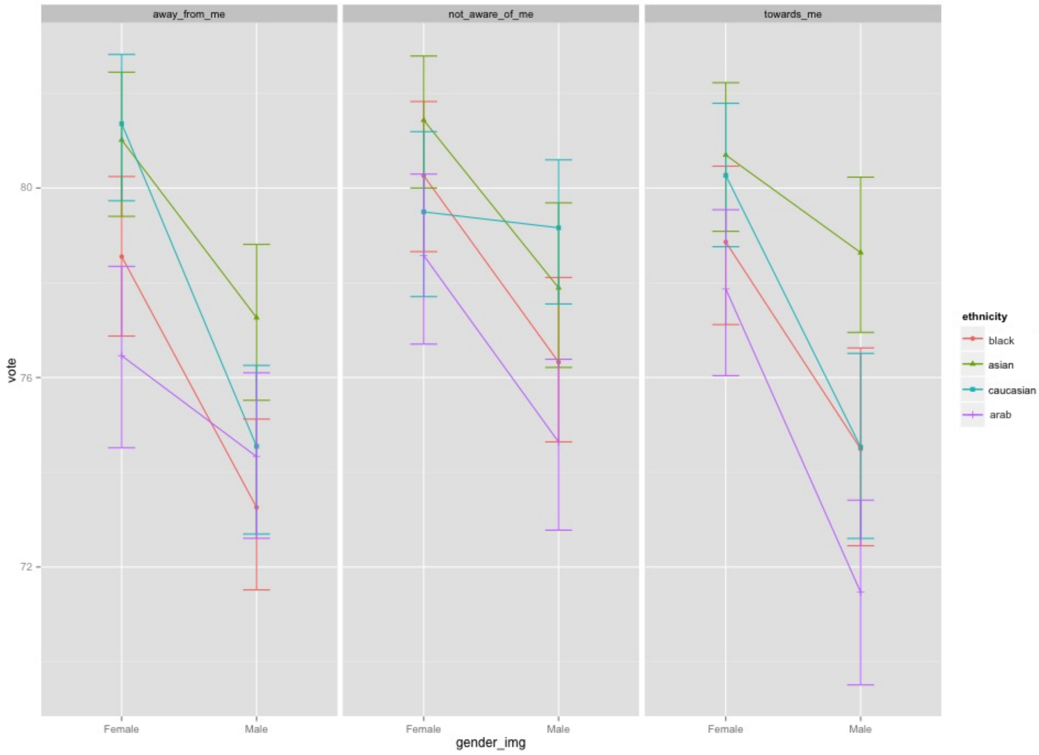
We observe that generally grown–up (green) and elder (red) people follow a similar pattern, which is that elder people are perceived as safer than grown–ups – especially when they looked towards the user. Teenagers (blue) show biggest differences between each facing–direction: while being perceived as least safe when facing away from the user, they are perceived as overall safest when not being aware.

$Gi : Fi$ – I found also a significant interaction between *gender* and the *facing–direction* of the person in the image ($F(2, 7885) = 3.77, p < 0.01$). Both, men and women show a similar pattern, which is that they are both perceived as safest when not being aware, and least safe when facing towards the user. However, I observe that generally facing–directions show bigger effects for men (towards ($p < 0.001$), not aware ($p < 0.05$), away ($p < 0.01$)) compared to women (towards ($p < 0.05$), not aware ($p < 0.05$), away ($p < 0.05$)). Furthermore, while there are minor differences in safety perception between looking at the user or looking away from him/her, I observe major differences when they are not aware of the user: while both, men and women, are perceived as safest in this situation, facing away shows a bigger positive effect for men compared to women (see Figure 5.4(e)).

While facing–towards and facing–away from the user shows similar outcome for men and women, we can clearly see the increase of difference for men when they are not aware.

$Gi : Ei : Fi$ – Besides two–way–interactions, I also found a significant three–way–interaction between *gender*, *ethnicity* and the *facing–direction* of the person in the image ($F(6, 7885) = 2.97, p < 0.01$).

When looking away from the user, I found differences in perception, depending on gender and ethnicity: While situations portraying women of Caucasian ethnicity ($p < 0.05$) were found to be perceived as safest, situations portraying men of Asian ethnicity ($p < 0.01$) were found to be perceived as safest. In addition, I found that while situations portraying women of Arab ethnicity ($p < 0.05$) were found to be perceived as least safe, situations portraying men of Black ethnicity ($p < 0.05$) were



(a) **(b)** **(c)**

Figure 5.5: Three–way Interaction plot, showing the interaction between *gender*, *ethnicity* and the *facing–direction* of the person in the image and resulting safety perception (from left to right): (a) facing away from the user, (b) not aware of the user and (c) towards the user. Y–axis shows means of perceived safety scores, x–axis shows the gender and colours indicate ethnicity of each interaction.

found to be perceived as least safe when looking away from the user (in Figure 5.5(a)).

When not being aware of the user, I also found differences in perception, depending on people's gender and ethnicity: For both genders, I found situations portraying Arab men ($p < 0.05$) and women ($p < 0.05$) to be perceived as least safe when not being aware of the user. While Caucasian men were perceived as safest ($p < 0.01$), Caucasian women were perceived only as second least safest ($p < 0.01$) when not being aware of the user. Additionally, while for situations portraying all other ethnicities and for all three facing–directions I observe that men are always perceived as significantly less safe compared to women, Caucasian men and women are perceived in a very similar way when they are not aware ($p < 0.05$) (in Figure 5.5(b)).

Furthermore, when facing towards the user, I also found differences in perception, depending on people's gender and ethnicity: Caucasian women ($p < 0.01$) were found to be perceived as second safest, followed by situations portraying Black women ($p <$

0.05). For men I observe a similar pattern, but with hardly any difference between situations portraying the two ethnicities (Caucasian ($p < 0.001$), Black ($p < 0.001$)), suggesting that these groups of men are perceived in the same way in terms of safety (in Figure 5.5(c)).

In Figure 5.5 I visualize these three–way interactions, depending on the *facing–direction* from left to right: looking away, not being aware and looking towards the user. While facing–towards and facing–away from the user show similar outcome in all ethnicities except for Caucasian for both men and women, we observe differences in perception for Caucasian and Arab people when looking away, depending on their gender.

These examples show clearly how differences within demographic groups of people we see matters to our safety perception. I now turn my attention to demographics characteristics of crowdworkers.

5.3.2 Who we are

Safety perceptions not only depend on *who we see* but also on *who we are*, the demographics and geographical background of the crowdworker. To analyse safety perceptions based on who the crowdworkers are, I divided my dataset in a next step according to the background information provided by the user: *gender, age, ethnicity, from London, where in London, continent*. As mentioned above, I received very diverse sample sizes for each of the properties, leading to limitations in terms of further evaluations due to lack of statistical power for smaller groups, such as Asian people or 0-20 year olds. From my data, I was able to reach statistical power for *gender, age* and *from London only*.

Table 5.4 shows the main effects related to the different crowdworker groups, related to age and location, whether he/she is from London of not. Next I will summarize these findings. Compared to the overall findings presented in Table 5.2, where responses from all crowdworkers were aggregated, I now focus on the differences in perceptions depending on crowdworker's background.

Gi *– Gender of the person in the image*: for all crowdworker groups, I found that the gender of the person in the image matters in a similar way. That is, men are

User Groups	*var name*	*contained Properties*	*Mean M*	
0 – 20	Gender image Gi	male	67.6	**
		female	73.13	**
	Ethnicity image Ei	Asian	73.37	*
		Arab	66.65	*
21-40	Gender image Gi	male	75.61	**
		female	79.75	**
	Ethnicity image Ei	Asian	79.84	***
		Black	77.16	**
		Caucasian	78.33	**
		Arab	75.48	***
	Facing image Fi	away from me	77.01	**
		not aware of me	78.77	**
		towards me	77.29	***
41+	Gender image Gi	male	76.7	**
		female	80.22	**
from London	Gender image Gi	male	76.0	***
		female	80.62	***
	Ethnicity image Ei	Asian	80.22	**
		Black	77.53	*
		Caucasian	79.06	*
		Arab	76.44	*
not from London	Gender image Gi	male	75.21	***
		female	78.96	***
	Ethnicity image Ei	Asian	79.08	***
		Black	76.61	**
		Caucasian	77.67	*
		Arab	75.06	***
	Facing image Fi	away from me	76.61	***
		not aware of me	78.11	**
		towards me	76.55	**

Table 5.4: Table showing breakdown of significant demographic variables, with each contained properties and their means of safety scores, ranging from 0 (indicating not safe) to 100 (indicating very safe), p<0.001 '***' p<0.01 '**' p<0.05 '*'.

perceived less safe than women. When I analysed answers by the age group of the respondents, I found significant differences for rating means (M) between users aged between 21-40 ($F(1, 5838) = 97.92, p < 0.001$), compared to users between 0-20 ($F(1, 267) = 11.93, p < 0.001$) and users 41+ ($F(1, 1636) = 21.24, p < 0.001$). I also observe differences when distinguishing crowdworkers from London ($F(1, 2878) = 62.18, p < 0.001$), and from outside London ($F(1, 4942) = 67.81, p < 0.001$) and found for both groups a similar significance ($p < 0.001$).

Ei – *Ethnicity of the person in the image*: the ethnic background of the person in the image was found to have significant effect on all groups, except crowdworkers who were older than 41. For crowdworkers who were younger than 21 ($F(3, 267) = 3.13, p < 0.01$) post hoc analysis revealed that only Asian ($p < 0.05$) and Arab people ($p < 0.05$) had significant effect on the received safety scores, suggesting that Asian people were perceived as most, Arab people as least safe. For the remaining groups of crowdworkers, post hoc analysis shows significant findings for all ethnicities in the image and in a similar pattern: users in the age group 21-40 ($F(3, 5838) = 19.39, p < 0.001$) perceived situations portraying Asian people as most safe ($p < 0.001$), followed by situations portraying Caucasian ($p < 0.01$), Black ($p < 0.01$) and Arab people ($p < 0.001$). Users from London ($F(3, 2871) = 8.03, p < 0.001$) perceived situations portraying Asian people as most safe ($p < 0.01$), followed by situations portraying Caucasian ($p < 0.05$), Black ($p < 0.05$) and Arab people ($p < 0.05$). And users from outside London ($F(3, 4942) = 13.87, p < 0.001$) perceived situations portraying Asian people as most safe ($p < 0.001$), followed by situations portraying Caucasian ($p < 0.05$), Black ($p < 0.01$) and Arab ($p < 0.001$) people. .

Fi – *Facing–direction of the person in the image*: facing–direction of the person shown in the image was affecting users in the age group 21-40 years ($F(2, 5838) = 6.84, p < 0.001$) or users that were not from London ($F(2, 4942) = 5.05, p < 0.001$). Users in the age group 21-40 years perceived people who are not aware of them as safest ($p < 0.01$), followed by people who look towards them ($p < 0.001$) and people that look away from them ($p < 0.01$) as least safest. Users who were not from London perceived people looking towards them as least safe ($p < 0.01$), followed by people who look away from them ($p < 0.001$) and people that are not aware of them ($p < 0.01$) as safest.

In summary, crowdsourced safety perceptions differ between each discussed group of crowdworkers, as different factors influence different people. For the age of crowdworkers, I found all three groups sharing the same opinion about *gender* (i.e., men trigger a lower feeling of safety compared to women). I also found that *ethnicity* matters more to both younger and grown–up people: while opinion about images depicting Asian and Arabs were found to be similar between the two groups, grown–ups showed larger differences in their ratings of images depicting Caucasian and Black people. *Ethnicity* was less important to the elder crowdworkers. *Facing–direction* was found to have only an effect on grown–ups. For the variable *from London or not*, I found that both groups share a similar opinion about gender and ethnicity of the people in the image, but *facing–direction* mattered only to people from outside London.

5.4 Discussion

5.4.1 Summary

In this chapter I have explored the viability of using online crowdsourcing to gather safety perceptions about people, and using the collected data to validate theories quantitatively. In doing so, I was able to support a number of established theories on large scale, previously only studied qualitatively (Zako, 2009; Felson and Clarke, 1998; Day, 1999; Pain, 2001). I have also been able to explore new factors, such as the *facing–direction* and *concealment* of a person that provide a more nuanced understanding of people's perceptions of risk. Furthermore, interactions between such features can now be easily analysed, as I have shown on three examples, thus giving power to researchers in social and urban studies to identify more features and in more depth.

5.4.2 Limitations

In terms of representativeness of answers, I hit a barrier, as crowdsourcing inevitably has self–selection bias. In studies where it matters who the respondents are, such as for safety perception towards people, I show up limitations of the method. By not being able to control who makes up the crowd, certain demographics of crowdworkers were not reached. This is an important limitation that researchers aiming to use crowdwork to validate theories need to be aware of. In fact, crowdsourcing has shown to suffer from self–selection bias due to lack of crowd–control in various fields. For instance, Quattrone et al. (2015) discuss bias found in spatial crowdsourcing datasets in the case of Open Street Map (OSM) and found that there is significant geographic bias. Kazai et al. (2012) showed that lack of crowd–control has an effect on the quality of task outcome with significant quality differences for a number of experimental tasks between Asian and American crowdworkers. Online market–places, such as Amazon Mechanical Turk (AMT), aim to offer control of crowd selection, but in fact show homogenity of their population as well, such as education level and nationality in worker demographics. For instance, Ross et al. (2010) showed that demographics of AMT workers have simply been shifting from moderate–income U.S. citizens to young educated people from India. Furthermore, Ipeirotis (2010) found that turkers were younger, mainly female, with low income and live in smaller families compared to U.S. internet users.

In the case of urban theories of safety, perceptions vary between different groups, hence it is important for the study designer being able to control variables of the crowd to make sure to include relevant demographics. If theories require specific demographics or compositions, then either we detect who is under–representated and use alternative methods, such as interviews, or we find ways of reaching out to them. There exist different approaches to engage and motivate specific crowds, based on gamification for instance (von Ahl and Dabbish, 2008). Still, it is not clear *who* they are attracting. In my studies, groups I was not able to reach in large enough numbers include certain ethnicities, and from specific geographic locations.

Another limitation of the study was the more or less arbitrary choice of age categories I have defined participants by, since I could not find a clear definition of age–groups (as for instance, defined by (ONS, 2015)). This leads to limitations of my results, as the group of 41+ year olds does not allow me adequately to distinguish anyone over the age of 40, and hence results of my study might not reflect clear age–related differences.

Furthermore, running a study that is based on images that have been selected by the researcher always includes a risk of selection bias based on researchers background, which might affect results. With the aim to minimize this effect, I have used broader image samples, but not being cross–validated through others, results still might have been affected. It is on the reseracher to avoid such limitations by cross–validation by other researchers from other backgrounds.

5.4.3 Implications

The method I have proposed has both theoretical and practical implications. From a theoretical standpoint, the method offers social science researchers a new way to investigate past theories on human perception based on their appearance, as well as develop new ones. Urban designers can include findings of this work (and follow–ups) and put them in an urban context, using this method, to evaluate how presence of different people demographics in urban space might affect our perception in the urban environment. Thereby the method can be easily reapplied over time, with images covering different aspects of research.

From a practical standpoint, findings of these studies can inform urban design approaches, respecting the built environment *and* the people inhabiting it. Findings can be used for urban interventions, supporting urban walkability, diversity and hence less motor traffic.

5.4.4 Conclusion

In this chapter I have shown to what extent online crowdsourcing can be used when it matters *who* makes up the crowd, for the case of collecting safety perceptions about people, and pointed out an important limitation – the lack of crowd control – researchers have to be aware of.

However, in a city we never see other people in isolation, but in interaction with other properties, such as the built environment. It is this combination of people and space that Tuan (2001) defines as *place*. In the next chapter I will study this relationship.

Chapter 6

A Crowdsourcing Approach to Evaluate and Develop Fear of Crime Theories of the Urban Environment

Part of the work presented in this chapter also appeared as a full paper accepted to the 2nd International Conference on IoT in Urban Space (urb-iot 2016) 2016. [Acceptance Rate: 20%]:

Traunmueller, M., Marshall, P. and Capra, L., "...when you're a Stranger": Evaluating Safety Perceptions of (un)familiar Urban Places, *In Proceedings of the 2nd International Conference on IoT in Urban Space (urb-iot 2016)*, May 2016.

6.1 Introduction

In Chapter 2 I have reviewed work that evaluated people's safety perception towards the static *built environment* following a crowdsourcing approach using Google Street View images (Quercia et al., 2014; Salesses et al., 2013). However, by using Google Street View images, reviewed studies focus primarily on the built environment as this source barely shows people in the images. In doing so, they keep out *people dynamics*. Reviewed qualitative studies in Chapter 2 found that these *people dynamics*, such as demographic background of people (Zako, 2009; Felson and Clarke, 1998; Day, 1999; Pain, 2001) do matter to safety perception, and I have shown in the previous chapter how online images can be used to crowdsource these safety perceptions not only for the built environment, but for people as well.

In summary, both *people dynamics* and the *built environment* were found to affect people's safety perception when being discussed each on its own.

However, in cities we do not experience these two properties as separate entities, but in interaction with each other. People change the appearance of urban space throughout time and give it different meaning (Gehl, 2010; Tuan, 2001). While using images of static facades only show parts of the visual urban experience, especially in times of high gentrification all over western cities (Skogan, 1986). In fact, the formerly run down neighbourhood of Greenpoint in Brooklyn/NYC which has experienced a major social shift through gentrification throughout the last five years, was found to be misinterpreted by most participants of the PlacePulse study (Salesses et al., 2013). While population shifts rather quickly within a city, the built environment takes more time to adapt.

Therefore it is necessary to include both *static* and *dynamic* properties into the methodology of research informing each other, to understand fear of crime in an urban environment. I will now study visual properties from both, the *built environment* and the *people* inhabiting it, interacting with each other quantitatively and at scale, to advance theoretical understanding.

To do so, I follow a similar approach to the one used in Chapter 5. First, I select images of people, according to their safety perception score based on my findings from Chapter 5; I overlay them on Google Street View images, selected by their safety score defined by the *Streetscore* algorithm (Naik et al., 2014). I use online crowdsourcing to collect safety ratings, that I then analyze using analysis of covariance (ANCOVA).

Besides visual properties of built environment and people, literature shows that familiarity with a place (Milgram, 1977; Paulos and Goodman, 2004) affects how people perceive it in terms of safety. Therefore I will include familiarity as additional variable into the study. Furthermore, to test to what extent online crowdsourcing can be used to gather also survey–like comments at scale, I will include a commentbox allowing participants to give reasons for their voting decisions.

The remainder of this chapter is structured as follows: next I outline my method to crowdsource safety perceptions of urban places, including *built environment* and *people* for the case study of Greater London. Then I will define reserach question and describe my analysis steps. I then present the results of my study and discuss its limitations.

6.2 Method

In this section, I describe the method I propose to quantitatively evaluate fear of crime towards people in the urban environment. I start with identifying potential limitations of this approach I need to be aware of to interpret results. I then discuss the selection of images of the built enviornment and of people, and continue with the study preparation process and outline the analysis steps.

6.2.1 The challenges of using actively collected data for studying fear of crime in the city

As I have discussed in the previous chapter, the use of images for online crowdsourcing includes a number of limitations and risks I have to be aware of when interpreting the results. As we follow a similar method in this study, I want to outline upfront such limitations I expect impacting my results.

In this study, I validate the effect of people presence on our safety perception when walking throught the city. Therefore it is necessary to define a selection of urban backgrounds, based on their perceived safety. To do that I use the results from Streetscore (Naik et al., 2014), a computer vision algorithm that has been trained on Google Street View images and their perceived safety towards people, on a crowdsourcing platform. The algorithm was evaluated in terms of accuracy and showed a regession performance of 54% accuracy and a classification performance of 78%. Although promising, these results still contain a high degree of inaccuracy and might lead to flaws in my findings.

In particular, relying on the outcome of a crowdsourcing study, I can expect the algorithm to be biased in terms of participant demographics, as already discussed in the last chapter. Furthermore, the algorithm was trained on U.S. cities and such cities have differences in architecture, urban layout and features compared to European cities, such as London. To reduce flaws resulting from urban design properties, I conducted a pilot study on a selection of images, as further discussed in 6.2.2. .

In terms of pictures of people to overlay to the selected pictures of places, I used images from my previous study. To reduce the risk of selecting people with unique looks, I used a broader sample for each defined type of people. However, there might be properties that have been overseen, leading to flaws for the results. Furthermore, as

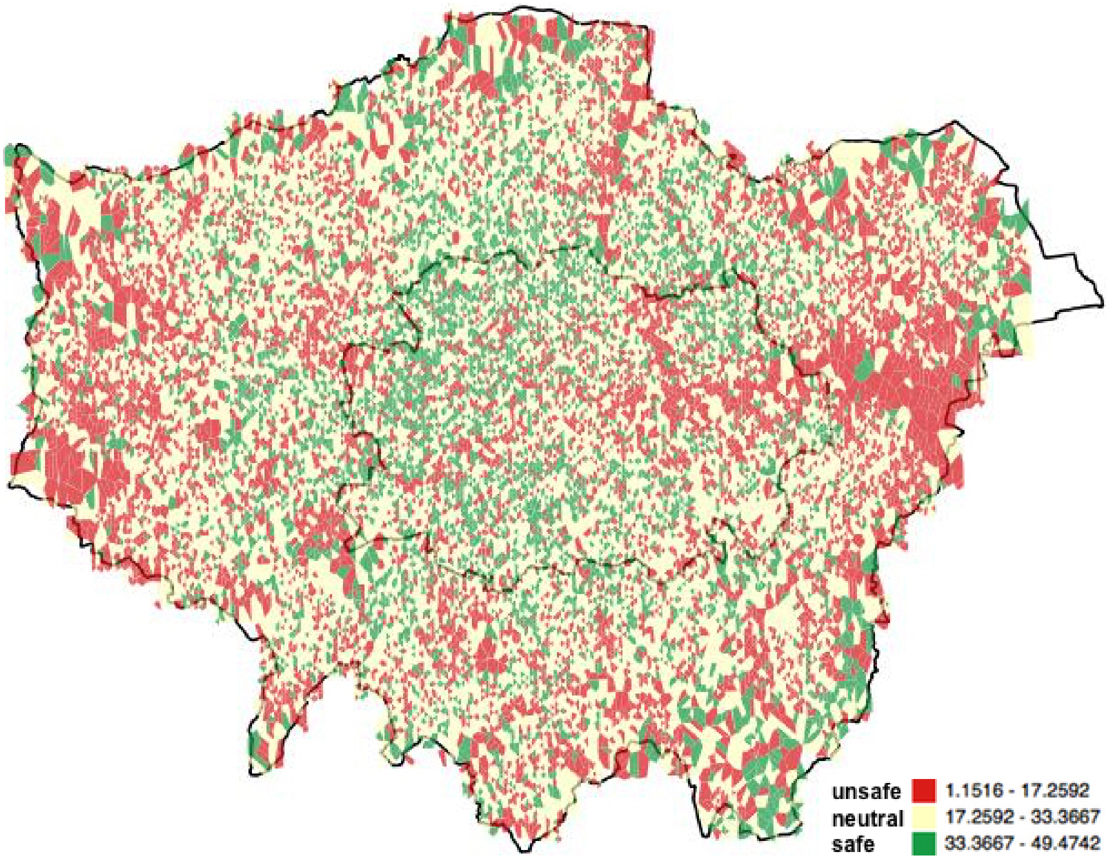


Figure 6.1: *Streetscore* map for Greater London, showing safety perceptions based on Google Street View images. Green areas are perceived as safe, yellow as neutral and red areas as unsafe.

outlined in the last chapter, as far as the image selection is based on the researcher they might be biased based on researcher's background.

As in the previous study, I was only able to reach a limited group of participant demographics. In this sense, outcome can not be generalized and has to be accounted specifically for this group. Result might differ when asking other demographics that my method was not able to reach. Being aware of these limitations, my aim is to show how to apply the method, as if the data would not suffer from these flaws.

6.2.2 Preparing images

To begin with, I created a pool of images showing various combinations of built environments and people. 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 I describe the selection process for each of them.

Selecting environments.

To study safety perceptions of urban places, I first identified a preselection of safe, unsafe or neutral perceived areas based on Google Street View images, as shown in Figure 6.1. To do so, I ran the *Streetscore* algorithm for the Greater London area. Based on previous work (Salesses et al., 2013), *Streetscore* is a computer vision algorithm that aims to define Google Street View images according to human safety perception based on the composition of the image (Naik et al., 2014). This step is necessary to evaluate changes in safety perception due to the presence of people overlaid on the image. To do so, I selected 30 images, including 10 safe, 10 neutral and 10 unsafe perceived environments: According to their scores, I selected 10 images from first (safe), 10 images from second/third (neutral) and 10 images from fourth (unsafe) quartile. However, as the *Streetscore* algrithm was trained on U.S. American cities, it was not sure if it would work well for a European city, such as London. Therefore I ran a small pilot study where I asked people to rate their perceived safety of these 30 images.

Method. I 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. I ran the study for two weeks in August 2015, advertised 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. Results of the pilot study showed a strong user agreement about perception of the environment ($F(2, 12) = 26.98, p < 0.001$) and confirmed the *Streetscore* classification for my image selection. More specifically, I found significant differences in mean scores (mS) between safe ($mS = 60.5$) and neutral ($mS = 52.4$) images ($p < 0.05$), unsafe ($mS = 43.8$) and neutral images ($p < 0.01$), and unsafe and safe images ($p < 0.001$), as shown in Figure 6.2.

Discussion. These findings show that similar properties of the built environment, such as the colour of green or differences in scale, affect people in the same way, no matter if they are in Europe or the U.S. This might result from a similar kind of cultural background people from the western world share, that informs their perceptions. On the other side, these findings also show that cities of the western world, besides all their differences in urban design, still seem to be not too different after all.

In the end, I selected for each attribute of safe, neutral and unsafe 5 distinct Google

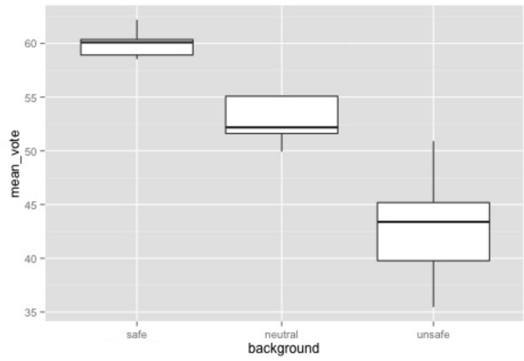


Figure 6.2: Online pilot study boxplot, showing the mean safety scores for safe, neutral and unsafe environments (from left to right).

Street View images. More precisely, I selected images with the smallest standard deviation of their mean safety scores quartiles: for safe from first quartile, for neutral from second and third quartile, and for unsafe from fourth quartile. In doing so, I ensured that selected images would be perceived in the same way by the majority of participants.

Safety perceptions depend on *who we are*, not just on what we see. Crowdsourcing often fails to reach broad demographics (Sen et al., 2015). To ensure representativeness of our selected images of built environments for different demographics, I conducted in addition a qualitative study with selected 15 images following a speak aloud approach (Ericsson and Simon, 1980).

Method. I interviewed 13 people living in London, between the age of 21-52 years, including 8 female and 5 male. With the aim to grasp people's first impression in terms of their safety perception, I kept myself in the background and encouraged them to speak open and freely about their general opinion and perception for each image. Their ethnicities included Caucasian, Black, 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.

Results. After transcribing audio recordings and safety votes, my data showed similar results to the findings of the quantitative part, such as that images including white houses or a high ratio of green colour were perceived from the majority as safe (86%), while images with big walls or narrow alleys were mostly perceived (93%) as dangerous. In Figure 6.3 we show three examples for each type of environment.

(a) safe perceived environments

(b) neutral perceived environments

(c) unsafe perceived environments

Figure 6.3: Environments: Image examples for safe (a.), neutral (b.) and unsafe (c.) rated environments as used in the pilot study (from top to bottom). Location map and road names were blurred out to ensure place anonymity.

Selecting people.

Having defined a selection of urban backgrounds, I next selected images of people to be overlaid on the top of them. I used the findings of Chapter 5 to define sets of safe, unsafe or neutral perceived types of people. Therefore I selected the 15 images in terms of mean safety votes and their standard deviations: 5 for each attribute of safe (first quartile), neutral (second and third quartile) and unsafe (fourth quartile). All images used in the pilot study were selected from online repositories under a creative commons license.

Overall, people who are perceived as unsafe included mostly younger men, while safe perceived people mostly included female and elder people. Neutral people included mostly young and elder men.

Overlaying people on backgrounds.

To study specifically the interaction between *people* in the *built environment* in relation to safety perceptions, it is necessary to combine my two selections of background im-

	safe Person	*unsafe Person*
safe Environment		
unsafe Environment		

Table 6.1: 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.

ages and images of people. Now I was able to do so: I created various combinations overlaying each of them with one another, as shown in Table 6.1. 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 (Hall, 1966). I used every image of a person on every image of background, resulting in a total set of 225 images to be used in my study.

6.2.3 Data Collection via Streetwise

With my images at hand, I was now able to present them to an audience to crowdsource safety perception data. To evaluate fear of crime theories according to familiarity and visual properties of built environment and people, I had to modify my crowdsourcing platform as presented in Chapter 5. Besides asking the audience about their demographic details and their safety vote, *Streetwise* asks them also to vote for familiarity with a situation; and they were also invited to comment about their vote via textbox (see Figure 6.4).

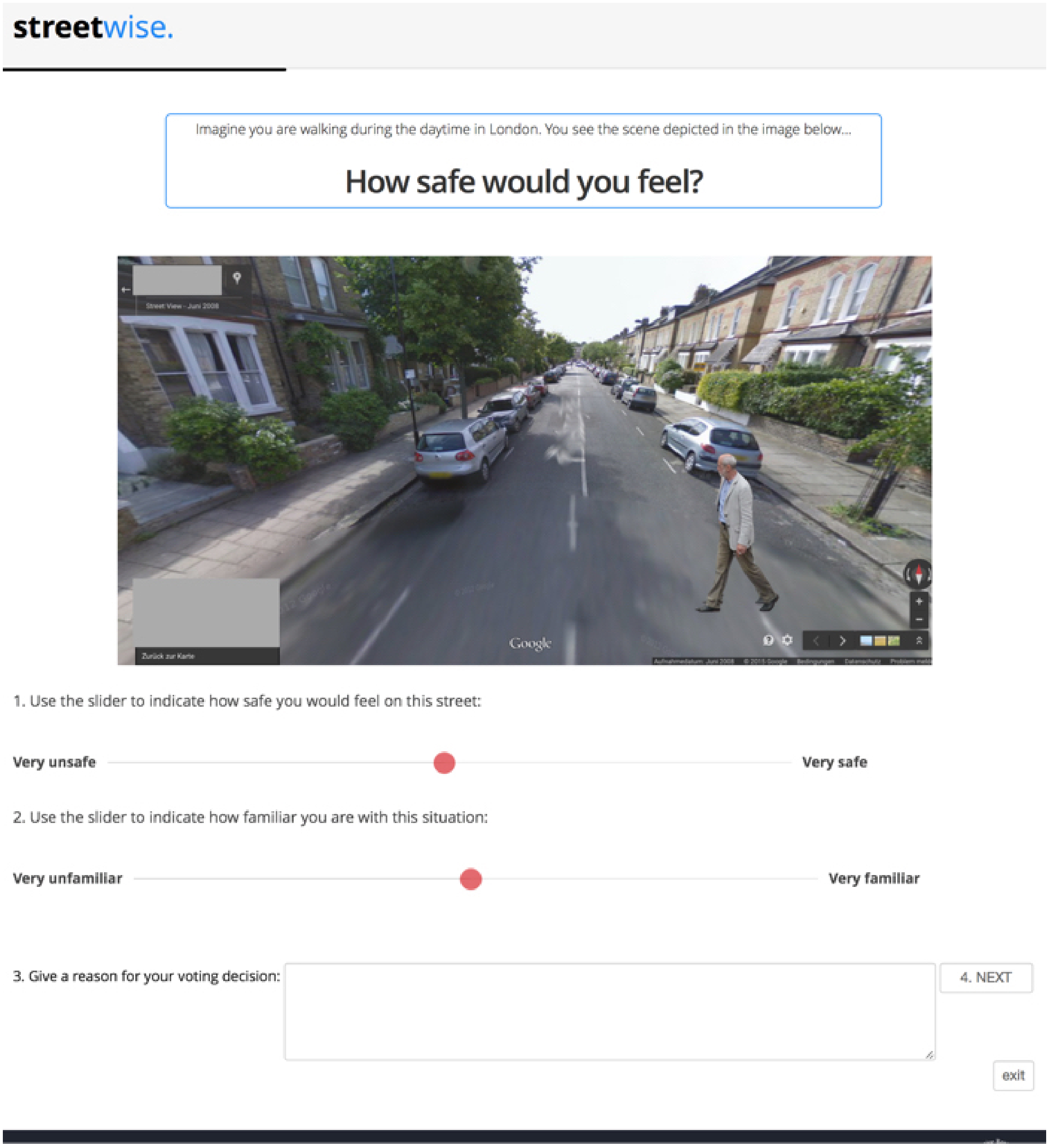


Figure 6.4: Web–based user–interface of *Streetwise*, showing one image of a person in an urban environment at a time. The user is asked to rate his/her perception of safety and familiarity with the situation according to the image on the sliders, and to give reasons. By pressing the 'Next' button a new image is presented; by pressing 'Exit' the study can be ended at any time.

I 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 with the situation depicted in the image (1 = very unfamiliar, 100 = very familiar). In addition, I asked the user to comment on the voting decision in a commenting box.

User comments, in connection with the vote, gave us a good indication about

Figure 6.5: Image examples, showing situations that were expected to raise the fear level of the user, due to showing dangerous situation.

seriousness of user feedback and hence enabled us to detect meaningless data I could exclude from the study. I also collected the time–stamp for each click, indicating how long the user thought about the vote before submitting. Furthermore, I 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, as seen in Figure 6.5. Including these images to my image pool, I 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.

I obtained ethical approval to conduct the study in September 2015. I 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). Users could share the webpage link via embedded Facebook–Like and Tweet buttons. In parallel, I used Amazon Mechanical Turk (AMT) to recruit crowdworkers, to speed up the data collection process.

Throughout these two months, I was 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 to make a serious judgement, as indicated by the time–stamp (Willis and Todorov, 2006)) from both, I 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, I found a high majority of Caucasian (78%), followed by

Black (8%) and Asian (6%) people. I merged the two datasets as received from AMT crowdworkers and participants recruited over social media and normalized safety and familiarity scores to a range between 0 and 1.

6.2.4 Research Question

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

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

6.2.5 Analysis

Familiarity with a situation is expected to impact people's safety perception (Milgram, 1977; Paulos and Goodman, 2004). As I could not control this variable, I included familiarity as covariate besides the categorical variables of type of built environment and type of person in the image. To study them I used an Analysis of Covariance (ANCOVA) with planned contrasts and post hoc analysis, using the Tukey correction (Tukey, 1949) for multiple comparisons to detect significant effects. The planned contrasts take a neutral perceived type of person and a neutral perceived environment as baseline to compare against other factors. Looking at the frequency distribution of ratings I 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. As done in Chapter 5, I used the *Aligned Rank Transform* (Wobbrock et al., 2011) for non–parametric factorial data analysis, which aligns the data in a pre–processing step before applying averaged ranks, to transform my data so that I could build an ANCOVA model. With transformed data, I was then able to conduct my analysis.

6.3 Results

I started studying people's perception of safety in the urban environment by analysing main effects. Therefore I took the safety score as the dependent variable, variables of built environment and the person – if perceived as safe, neutral or unsafe – in the image as independent variables and familiarity–score as covariate. In Table 6.2 I present the F–scores and p–values for each variable of the model as main effect and in interaction

Effect	*var name*	Variable	*F–Score*		*contained Factors*	*Mean*
Main effect	Fi	Familiarity	2614.91	***	covariate	
	Bi	Built Environment	141.12	***	safe	78.7
					neutral	70.8
					unsafe	59.1
	Pi	Person	112.04	***	safe	74.6
					neutral	73.5
					unsafe	60.0
Interaction	$Pi : Bi$	Built Environment x Person	2.7			

Table 6.2: Table showing main effects and interactions with their F–score, with p–value: 0.001 '***' 0.01 '**' 0.05 '*'.

with eachother. Next I summarize the findings.

Main effects. The covariate, familiarity with the situation depicted in the image Fi, was significantly related to the user's perception of safety ($F(1, 3853) = 2614.91, p < 0.001$). There was also a significant effect of the built environment Bi ($F(2, 3853) = 141.12, p < 0.001$) and people Pi ($F(2, 3853) = 112.04, p < 0.001$) on safety perception after controlling for the effect of familiarity.

These results indicate that familiarity with a place Fi is the single most important variable that affects our safety perception, compared to visual properties of the environment. To ensure that this is the case, I next computed effect sizes for each of our variables and related results to Cohen's rules of thumb (Cohen, 1988), approving these indications: The covariate familiarity showed a large effect size with Fi r = 0.47, built environment a medium effect size with Bi η^2 = 0.04 and people also a medium effect size with Pi η^2 = 0.03.

For discussed categorical variables (type of built environment Bi, type of person in the image Pi) planned contrasts revealed no surprises: looking at the mean safety scores as described in Table 6.2, safe environments were perceived as significantly safer than neutral ($p < 0.001$) or unsafe environments ($p < 0.001$), and neutral environments were perceived as significantly safer compared to unsafe environments ($p < 0.01$). Safe people were perceived as significantly safer compared to unsafe people ($p < 0.001$), while neutral people were found not to be significantly different to safe people. Neutral people were found to be significantly safer compared to unsafe people ($p < 0.05$).

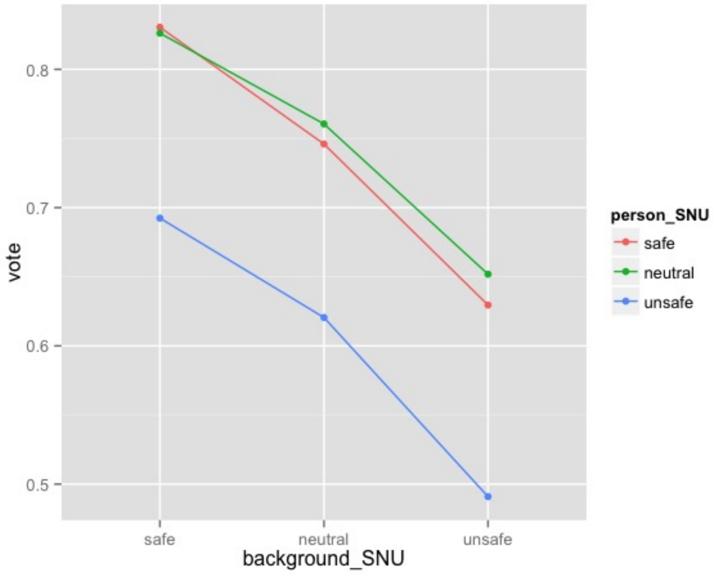


Figure 6.6: Interaction plot between safe, neutral and unsafe perceived built environments and people. y–axis shows means of perceived safety scores, x–axis shows the type of built environment and colours indicate the type of people.

Interactions. My results for the main effects show to what extent appearance of built environment Bi, appearance of people Pi and place familiarity Fi affect people's safety perception. However, we knew already from previous work (Quercia et al., 2014; Salesses et al., 2013) and the last chapter that people and the environment affect people's safety perception each on their own, these findings are not surprising. In reality, however, these factors are not isolated, but happen in interaction with each other, so next I explored their interaction.

$Pi : Bi$ – According to my own perception and preliminary discussions with colleagues prior to this study, I expected to find a significant interaction effect between the presence of people and the built environment. For instance, I expected that the presence of a safe person in an unsafe environment would affect safety perceptions significantly towards safe, and vice versa. However, my findings show that there is no significant interaction between the person and the built environment depicted in the image ($F(4, 3853) = 2.7$) . Also I only found a very small effect size (η^2 = 0.001), according to Cohen (1988).

Looking at the interaction plot in Figure 6.6, we observe a simple additive relationship between people Pi and built environment Bi affecting safety perception. While safety perceptions towards all three types of people – safe (green), neutral (red) and unsafe (blue) – follow a similar relationship, we observe that surprisingly not safe, but neutral people are perceived as safest in an unsafe environment. From that, we can only assume that individual properties of selected people in the images might have an effect on the outcome.

However, these results show clearly that my hypothesis, that there is an interaction effect of presence of different people in the environment to our safety perception, must be rejected.

6.3.1 Thematic Analysis.

Besides scores for safety perception and familiarity with the presented situation, we asked participants to comment on their voting decisions via a text box. With the inclusion of comments, we aimed to go beyond numerical scores to investigate more qualitatively about what influences the voting decisions. Answers were collected and stored as string to be analyzed qualitatively using inductive thematic analysis as described by Hayes (2000).

Analysis.

I first divided the 5452 collected comments into a *safe*, *neutral* and an *unsafe* perceived group according to their safety scores. The *safe* perceived group included images from the first quartile ($s = 1902$), the *neutral* perceived group included images from the second and third quartile ($s = 2275$), the *unsafe* perceived group included images from the fourth quartile ($u = 1275$). As only a small part (7.8 %) of the *neutral* perceived group was commented on, I focussed on *safe* and *unsafe* perceived groups. From each of the two I picked randomly 10 comments per iteration, to do thematic analyses on them and marked out words to identify meaningful phrases that were relevant to the topic. After 10 iterations and marking 100 random comments (50 % *safe*, 50 % *unsafe*) words kept repeating as I reached saturation and could not gain anything new from the data (Furniss et al., 2011) and had covered all important topics. Marked words dealing with the same topic were grouped together into analytical categories. I continued this process in a next step on additional 300 randomly selected comments for each group and systematically reviewed my data to see if the words fit into defined categories

	Themes	Categories	*safe*	*unsafe*
1	People Demographics	**Gender**	woman, women, lady, girl, she, her	man, men, guy, boy, he,his
		Age	old	young
		Ethnicity	White, Asian	Black, Middle–Eastern
2	Built Environment	**Openess**	open, windows wide, visible	narrow, closed–in, alley, remote, hidden
		Maintanance	nice, clean, looked–after, cared–of	sketchy, abandonded, trash, run–down
		Age of Buildings	new, modern	old
		Type of Buildings	residential, green	industrial, business
3	Familiarity		normal, typical, regular, familiar, known	unknown, strange weird, unfamiliar
4	Economic Factors	**Wealth**	well–off, upscale, affluent,	rough
		Class	upper-class	low–class
		Type of Cars	expensive	cheap
		Income	rich	poor, low–income
		Ownership	owner	rental, tenants, council
5	Environmental Conditions	**Time of Day**	day, daytime	night
		Brightness	bright, well–lit, lit–up, light	dark
		Weather	sunny	
6	Area Activity		people, busy, calm, quiet, peacefull, cars	lack of people, lack of cars, deserted, unpopulated, alone
7	Type of Area		residential	industrial

Table 6.3: Table showing defined themes and categories based on found "signal words" for *safe* (left) and *unsafe* (right) perceived images. Categories ranked by their importance: **bold** categories appeared most.

or if there are more to add. I noticed after this run, that the categories remained the same, but were still enriched by new words. To ensure that not only categories, but also words for each were not changing anymore, I analyzed another 100 random comments (50 % *safe*, 50 % *unsafe*) and came to a stable condition with no more words to add. Finally, I identified 15 categories which I summarized in seven themes, as shown in Table 6.3. Overall I have analyzed 1000 answers or 18.3 % of the total collected data.

Results.

Although participants were not obliged to use the comment box, I found the majority of answers to have a comment on their voting decision (68.7 %). On average, answers were 10.32 words long (SD = 11.30). This high willingness of participants to give extra incentives to their voting decisions shows the potential of this method to gather not only quantitative voting–data, but also survey–data on large scale. I was then able to reveal seven key themes from found "signal words", consisting of 15 categories, as shown in Table 6.3. Three themes were found to support my findings from the quantitative studies, such as *People Demographics* (row 1), *Built Environment* (row 2) and *Familiarity* (row 3). In addition to these, I identified four themes that have not been covered, such as *Economic Factors* (row 4), *Environmental Conditions* (row 5), *Area*

Activity (row 6) and *Type of Area* (row 7). I now elaborate on each of them.

- *People Demographics.* One theme that influences people's safety perception in the city is cues from visual properties of other people they encounter on the street, based on their demographic background. User comments show that especially people's *gender* was found to play a significant role. While the presence of women was found to be linked to safe rated situations (e.g., *"Judging from the lady, she feels safe walking in the street. So I'd feel quite safe as well."*), the presence of men was found to be linked to unsafe situations (e.g., *"This dude looks suspicious... maybe nothing to worry about, but I'd stay concious."*). Moreover, I found that the presence of men can decrease our safety perception significantly, even in safely perceived environments (e.g., *"Looks like a nice residential neighbourhood with the only concern the man standing by the car in the street. I would keep my eye on him. He looks suspicious"*).

  Another demographic property of people that was mentioned frequently was their *age*. According to my collected comments, the presence of young people has a negative effect on the overall safety perception in the city, as an area may *"appear very nice but the kid there makes me wonder if there are gangs or troublemakers in the area."*. Activity and facing direction of young people could increase this negative effect when just standing on the side of the road and looking at you, as this example shows: *"youth with hand in pocket looking directly at me... probably not a problem but i would not engage him in eye contact and would walk to opposite pavement."*. The negative effect was increased especially for young men: *"I wouldn't want to be in close proximity to a young male"*. At the same time, the presence of men was mostly perceived as safe when they were senior, such as these examples show: *"If the old man can walk safely here, I'd feel quite safe"*, *"I would feel safe in this very nice neighbourhood with a senior man walking around."* or *"I don't care for the warehouse type area but the senior man walking would probably make me feel more comfortable"*. These findings reveal an interaction between gender and age: While the presence of young men was found to decrease safety of others, the presence of elder men increased it.

  A third demographic property that was mentioned frequently was people's *ethnic* *background*, which was mostly mentioned in relation to unsafe perceived situations. Black, or *"shady black guys"* were mentioned besides Middle–Eastern ethnicities through religion, revealing participant's prejudice (*"This looks like a muslim area and they don't have a good repuation. If I was to walk down it without the required dress i could be in big trouble."* or *"I think this is some muslim neighbourhood, although it looks clean."*) and uncertainty (*"I don't know about muslim neighbourhoods."*). These comments might result from the demographic background of study participants, as most of my participants (72%) came from the U.S. or Western Europe and may show prejudice coming from these demographics. However, unraveling this prejudice through these comments allows us to explore ethnic prejudice further for future work.

- *Built Environment.* Another theme includes features based on the design of the built environment. A main feature thereby was the general *openness* of the area presented in the picture. Open or wide was related to safe perceived images (e.g., *"It's wide open with plenty of room to avoid any bad potential situation. And there appears to be plenty of places to call out for help if needed."*) and closed, narrow or hidden to unsafe images (e.g., *"The street is very narrow and I see no alley ways or breaks between the buildings. I'd have to run a block to make left or right if attacked."* or *"I would feel a bit unsafe as there are many places for people to hide and jump out from"*).

  Another feature was the general *maintanance* or tidyness of the buildings. While clean and tidy was strongly linked to safe perceived situations (e.g., *"The neighbourhood looks very well kept. It is brightly lit. The houses and buildings are well kept and in good condition."*), a lack of maintanance was often linked to unsafe perceived situations (e.g., *"This looks a little run–down and not very inviting. I would be a little concerned walking through here."*). This perception was amplified considering buildings *age*, where old buildings lead to feelings of a lack of safety and young or modern buildings to safety as these example shows: *"Pretty empty looking street with some older looking buildings. There is a lone woman walking down the street... but that doesn't make me feel much safer"*.

- *Familiarity.* Many participants mentioned the degree of familiarity with the situation in their comments, to impact their decisions. Situations that were commented with “familiar”, “typical”, “normal” or “usual” were found to be perceived as safe, such as this comment shows: *“I’ve been living in urban environments (Boston and NYC) my whole life so this is very familiar to me and thus very safe feeling”*. Situations commented with “not familiar” or “strange” were found to be perceived as unsafe, as for instance: *“A residential area that I am not familiar with. Hmm... wouldn’t feel too safe”*.

- *Economic Factors.* The economic factors that define the visual properties of an area were found to have great impact on how safe people feel in an urban environment. Participants identified features in the situation presented, that indicated the *wealth* or *class* of the area. Generally wealth (e.g., *“This street looks like it has very expensive housing on it. I feel safe on this street because of this.”*) and higher classes (e.g., *“This looks like a more upper class neighbourhood... I’d feel safe.”*) were linked to safety, while poverty (e.g., *“Walls are crumbling and dirty looking make me think the area is inhabited by people who don’t take pride in the area and might be poorer.”*) and lower classes (e.g., *“Looks like lower middle class neighbourhood. Brick wall, etc...”*) were linked to lack of safety.

  Besides the appearance of the houses, especially the *type of cars* was used as indicator for wealth of an area as these examples show: *“Nice buildings and cars. Low crime appearing neighbourhood because of it.”* or *“Appears to be an urban scene with low priced cars. I would feel uncomfortable with my family in a location like this”*.

- *Environmental Conditions.* By using Google Street View images in the study, I aimed to make sure similar weather and time conditions between the images, as Google claims to take images at the same time of day and in similar weather conditions. However, participants of the study detected differences between the images that had an impact on their safety perception, such as the *time of day*, *brightness* or *weather*. *“Day”* (e.g., *“It look pretty residential and it’s daytime”*) or *“bright”* (e.g., *“This area is brightly–lit. I’d feel safe.”*) were found to be mentioned mostly with safe perceived situations. The lack of light was mostly mentioned with unsafe perceived situations, such as *“The building along the street* *look a little beat up and there isn't a whole lot of light on the street"*.

- *Area Activity.* The overall activity in the area was another theme I detected from my received comments. Attributes describing an active, but still calm neighbourhood were found to be linked to safety, while attributes describing a total lack of activity were found to be linked to unsafe perceptions. For instance, participants commented *"just fine... shave a few points off the safe–meter for lack of human activity and overall bareness withough many exit strategies, but just fine."* or rated another situation as very safe, because *"it looks like its in an area that is densely populatioed and would have high amount of security"*. On the other side, people rated images as less safe because *"there aren't many people or cars around. This makes me feel like the area is secluded like an alley or something. The street looks cared for but I would be worried if I encountered another person walking down the path"*.

- *Type of Area.* In close relation to people activity, I also found that the type of the area impacts people's safety perception. While mostly residential areas were perceived as safe (e.g., *"Residential neighbourhood... no worries at all – I feel safe."*), industrial areas were perceived as unsafe. (e.g., *"Industrial area where there would not be residents. Hmm... the woman there doesn't make any difference to me"*).

In summary, we see how properties found in the urban environment affect respondent's safety perception differently according to their familiarity with an urban place. While the built environment is affecting their safety perception both positively and negatively in unfamiliar situations, I found that it only affects them negatively in familiar situations. Furthermore, I found that some demographic properties of people they encounter on the streets affect them in familiar and unfamiliar situations, others do not. For instance, while gender of people matters to our respondents most in neutral situations, people's age matters most in unfamiliar situations. People's ethnical background was found for respondents to impact depending on the grade of familiarity with a situation differently: while Black 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 and experience. As identified in Chapter 2, safety perceptions are very personal and depend not only on *who we see*, but also on *who we are*. It is important to point out that my findings apply to a specific group of study participants, covering only certain demographics as shown in the last chapter. Next I will discuss these findings and their limitations.

6.4 Discussion

6.4.1 Summary

In this chapter, I explored quantitatively people's safety perception resulting from visual properties of the built environment, other people and place–familiarity. To do so, I 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. In doing so we have also tested the method's potential to gather survey–like data at scale. I found that familiarity has the biggest effect on people's safety perception, compared to the visual properties of the built envionment and people. I also found that there is no significant interaction effect resulting from these variables, affecting people's safety perception. Furthermore, my approach also shows its usefulness to detect themes in crowdsourced comments that are worth studying more.

6.4.2 Limitations

However, using online images to crowdsource safety perceptions of the urban environment brings up a number of limitations.

Built Environment. The question arises whether the 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 (Quercia et al., 2015). That is, we experience a city not just through single images one by one, but through movement, developing a sense of place over time (Heft and Nasar, 2000). My 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, as for instance sound using VR (Park, 2008). Findings can be combined with this approach, resulting in a more complete and in–depth study on urban safety perceptions resulting from these features.

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 this study, I focussed on London; we do not know whether the outcome would differ when running a smilar study for another city and how. However, the study could be easily repeated focussing on other cities as the method allows it.

People. Based on my findings of the previous chapter (Chapter 5), I discussed in this chapter 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 (Matei et al., 2001). 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, I discussed the influence of one person at a time only, and I 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*. In Chapter 5) I found that crowdsourcing methods might lead to results that are biased by their crowd. Demographic data gathered from my study participants reflects these circumstances, as I received feedback mostly from Caucasian, middle–aged U.S. Americans. Therefore my results reflect the opinion of this crowdworker demographics only.

6.4.3 Implications

My 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 collect 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 (Nasar and Fisher, 1993; Warr and Ellison, 2000) leading to less social interaction among the population and more motor traffic within the city (Warr and Ellison, 2000). The method proposed in this chapter 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. The findings emerging from the use of such method can then be used in practice 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.

6.4.4 Conclusion

In this chapter, I have explored quantitatively how visual properties found in the urban environment (such as the built environment and people inhabiting it) and place–familiarity affect in interaction people's safety perception. In doing so, I extended previous work that followed a similar approach, but focussed on the static *built environment* (Quercia et al., 2014; Salesses et al., 2013) or on the demographic background of *people* only (Chapter 5) as separate entities.

While both were found to impact people's safety perception each on its own, they have never been evaluated in interaction with each other. As this is the way how they affect people's safety perception in the urban environment (Gehl, 2010; Tuan, 2001), it is important to discuss them in interaction. My work extends research in this domain by offering a method that gives answers to these questions.

Chapter 7

Discussion and Conclusion

In this chapter I summarize the contributions of this thesis (Section 7.1) and how they will benefit different communities, including computer science, social sciences and urban studies (Section 7.2). Then I critically evaluate them, pointing out limitations (Section 7.3), and define future directions of research (Section 7.4).

7.1 Summary of Contributions

This thesis proposed a quantitative, data–driven methodology to validate theories about urban crime and fear of crime across communities of computer science, urban studies and social / criminological research. I used passively collected mobile phone data and actively crowdsourced data on safety perception, based on visual properties of urban places. By analysing these data I was then able to support individual theories, and also evaluate their interactions.

7.1.1 Chapter 4: People dynamics and crime in the city

In Chapter 4 I have presented a method to use mobile phone data to validate quantitatvely urban crime theories. From the mobile phone data I extracted metrics that act as proxies for urban crime theories and correlated them with crime data, showing that it is now possible to quantitatively investigate urban crime theories at large geographic scale and frequent intervals. Findings of this chapter point out the significance of *people dynamics* in the urban environment relating to crime activities and their validity for a contemporary metropolis like London.

7.1.2 Chapter 5: People dynamics and fear of crime towards people

In Chapter 5 I have explored the viability of using online crowdsourcing to gather safety perceptions about people, and using the collected data to test theories quantitatively. In doing so, I was able to confirm a number of established theories quantitatively, previously only elaborated qualitatively. I have explored new features, such as the *facing–direction* and *concealment* of a person. Furthermore, interactions between these features can now be easily analysed, as I have shown in three examples, thus giving a new approach to researchers in social and urban studies to identify more features and in more depth.

7.1.3 Chapter 6: People dynamics and fear of crime in the city

In Chapter 6, I have used a similar methodology as presented in Chapter 5 to explore quantitatively the relationship between visual properties found in the urban environment, place–familiarity and the resulting perception of safety. I extended the methodology and showed opportunities to crowdsource survey–like data, such as comments, besides user–votings at scale. By using findings from prior work on the topic, discussing these visual properties as seperate entities, I discussed them in combination with each other as they appear when walking through the city. I found familiarity of the place to be the single most important factor of our perception of safety. I also found that there is no significant interaction effect on people's safety perception resulting from presence of people and the built environment.

7.2 An interdisciplinary perspective

Findings that result from this thesis benefit different stakeholders from various communities, including computational, urban, social and criminological studies:

Computational social scientists: The work presented in Chapter 4 offers a novel method available to computational social scientists that could support future studies on urban life. Social studies mostly use static census data of a city's population to describe dynamic processes, as mentioned in Chapter 2. A city like London is highly diverse in its social, demographical and ecological properties that change its face throughout its geographical spread and time. With an increasing population shift of people moving into cities, these properties change even more quickly than in the past. As population within a city changes over time, so do the boroughs they live, work and socialize in. Properties of such a steadily changing process are difficult to obtain using static sources as census data. By using data sources representing changing properties dynamically, as mobile phone data, my methodology offers new opportunities to guide such studies.

Work presented in Chapter 5 and 6 offers computational social scientists opportunities to gather so called 'soft data', as for instance if people feel safe or not in an urban environment, at large scale. In addition, I have also shown the method's potential to gather survey–like data at scale in Chapter 6. As such data has been difficult to collect quantitatively, the method can be used to harness this 'soft data' more easily, opening doors to computational social scientists to develop and evaluate theories at large scale.

Urban studies: As reviewed in Chapter 2, architectural and urban studies have attempted to describe the relationship between the built environment and crime (Wolfe and Mennis, 2012; Hillier and Sahbaz, 2009; Sahbaz and Hiller, 2007) and show that there is a strong relationship between the two. However, the very same built environment is appropriated and used by different people for different purposes and in different ways throughout the day, which has been difficult to describe. With the method presented in Chapter 4 such properties are now easy to include into such studies to discuss urban crime in relation to the built environment and the population inhabiting it. Besides crime, the method offers researchers also the opportunity to study other urban phenomena, such as gentrification, where static data sources, for example census data, are limited in what they are able to express.

My work presented in Chapter 5 and 6 offers opportunities to urban studies to understand urban environments and how they are perceived by their population. The way people perceive the environment has great impact on sustainability of a city and urban life quality of its population: if people avoid feared places, the city's walkability decreases (Nasar and Fisher, 1993; Warr and Ellison, 2000) leading to less social interaction among the population and more motor traffic within the city (Warr and Ellison, 2000). To make cities more sustainable and increase urban life quality of its population, it is therefore important to understand such factors. The findings emerging from the use of such methods can be used in practice 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 software 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.

Criminologists: The method I have proposed in Chapter 4 offers criminologists a new way to investigate past crime theories, as well as develop new ones. While I have presented the use of this method to explore established theories for the case of Greater London, the same method could be used for other cities to advance knowledge in terms of the contexts within which past theories hold. The method also offers researchers the ability to be reapplied over time, so to detect possible changes that call for scientists to refine past theories or develop new ones. In my case for London and for a single time period, I have shown that some theories do not hold across all boroughs, thus calling for deeper qualitative investigations in selected areas. I foresee the proposed quantitative method to be used in conjunction with qualitative methods, during alternate phases of theory development and evaluation.

In combination with qualitative approaches, the method can be used first to quick test and refine theories under development, and then to validate findings at scale in different geographic contexts and at different times, to understand under what conditions they hold. Having detected where findings hold and where not, qualitative investigations are again necessary to gain deeper insight, so the two approaches – qualitative and quantitative – are both to be used.

Software developers: Developers can use my findings from Chapter 4 to build tools on top of them, for the benefit of different stakeholders, such as citizens, administrators and city planners. For example, citizens may appreciate predictive crime tools they can use to decide what areas of a city to explore safely and which to avoid; administrators may use tools that highlight time variations in the model, to monitor the impact of processes such as urbanisation and gentrification on an area's dynamics and crime activity; and city planners may use tools that highlight crime model similarities and differences across city neighbourhoods.

7.3 Limitations

In my work I show that data mining CDR (Call Detail Record) and crowdsourcing safety perceptions of urban places can help to unravel complex *people dynamics* in an urban environment that matter to (fear of) crime. In doing so, I faced a number of challenges that limited the work and will inform future work using similar methods.

7.3.1 Mining Big Data Sets

As identified in Chapter 3, research that is based on mining big data sets as methodology is highly dependant on the level of granularity of the data provided. My work of Chapter 4 was limited by different temporal unit of analysis used in the two datasets provided (i.e., crime data was recorded on a monthly basis, while foot–counts were recorded on a hourly basis), requiring us to operate at the coarser level of granularity. In doing so, I had to manipulate the data in a way that prevented me from studying crime more in–depth, as it varies through the days. These circumstances also lead to issues of identifying when a crime actually took place. As reporting times for different crime types may differ (HMIC, 2014), this can affect the validity of the results.

Geographically, the mobile phone data set was defined by a grid and not the geo–location, that lead to a lack of geographical accuracy. In my specific case for London, a city that can change dramatically within just a few foot steps, this lack of accuracy becomes a problem when interpreting the results. Furthermore, using a grid instead of geo–locations lead to problems to differentiate if people are on the streets or in buildings.

While the crime data set included geo–locations of each data point, these coordinates only represented an approximate and not the exact location of a crime, according to the open source platform (PoliceUK, 2013b). These circumstances lead to geographical inaccuracy as the data does not allow the accurate pinpointing of crime locations. Furthermore, the crime data set provided information about reported crime, leaving out the dark number of unreported crimes (HMIC, 2014) that have been excluded from the study.

However, these limitations result from the quality of the datasets at hand and not from the method that has been proposed. This thesis shows how the method can be applied to evaluate and develop theories of urban crime quantitatively. As new, more accurate datasets become available, I believe the validity of the method withstands.

Accuracy often is closely related to availability of data. Mobile phone data, for instance, is usually very sensitive and confidential and therefore difficult to get as academic researcher. The mobile phone dataset as used in presented study was received due to a big data challenge and was provided pre–processed. I was not able to receive detailled information from the mobile phone provider about how the data has been processed and how different types of people (such as, residents, visitors and workers) have been identified, which might have caused flaws for the study's results.

Also, by using mobile phone data for research purposes I have to be aware that I capture only some parts of population – in my case 25% – but exlude others, such as people using another provider, PayAsYouGo options or those who do not use mobile phones at all. While qualitative approaches include these groups as well, it is a limitation of the methodology that it does not capture the whole population.

7.3.2 Online Crowdsourcing

Using online crowdsourcing, it was difficult to reach an even distribution of user demographics. As defined in Chapter 2, safety perceptions are very personal, defined not only by *what we see*, but also by *who we are* and strongly relate to people's background and demographics. By not being able to control who makes up the crowd, I was able in Chapter 5 to gather sufficient data only for the demographic group of Caucasian and middle–aged participants, while other demographics of crowdworkers were not reached. Therefore results only reflect opinions of Caucasian middle–aged participants, and can not be generalized. This is an important limitation that researchers aiming to use crowdwork to validate theories need to be aware of. In the case of urban theories of safety, perceptions vary between different groups, hence it is important for the study designer to be able to control variables of the crowd to make sure to include relevant demographics.

Furthermore, using images beared a risk of selection–bias from my own background, that affect results. In Chapter 5 I aimed to minimize this effect by using a broader sample.

Image selection in Chapter 6 was done with the help of Streetscore, a computer vision algorithm, to select Google Street View images according to their perceived safety. However, as long as Streetscore has been developed based on a crowdsourcing study, my results might suffer from biased work in terms of their participant demography.

Furthermore, the question arises whether people and three dimensional urban space can be represented on two dimensional images on a computer screen. Besides visual properties, people and urban space are being defined by many other variables that we perceive subconciously through other senses, such as people's behaviour in a situation or the sounds and smells of urban space (Quercia et al., 2015). We experience a city and its inhabitants not just through single static images one by one, but through movement, developing a sense of place over time (Heft and Nasar, 2000). My platforms based on images do not capture any of these, but focus on the visual characteristics of people and the urban environment only, and therefore have to be seen as a first step that leaves space for future research to add other properties to it.

Another limitation emerges from urban design questions, as 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 my study, I focussed on London; I do not know whether the outcome would differ when running a smilar study for another city and how. However, the study could be easily repeated focussing on other cities as the method allows it.

7.4 Future Directions

In the previous section I have identified important limitations of methods used in this thesis, that offer opportunities for future research to work on.

Mobile phone data and urban crime: I have shown in Chapter 4 how mining mobile phone data can be used to test established urban crime theories on a quantitative level, and even to predict crime in a city. Many of our limitations were – as described above – caused by poor data quality, while we have shown the potential of the method. As city agencies and businesses, such as mobile phone companies, increasingly collaborate with academic research with the aim to define future urban strategies (NYU, 2016), more accurate data will become accessible in the next years and for research less costly to get hands on. For here presented work this means opportunities to be continuously refined, expanded and improved in multiple ways.

As next step the model can be refined and expanded, so to incorporate properties of people dynamics, the built environment, and census within a single framework, to predict not only crime activity with greater accuracy, but also to understand the dependencies between all such variables in relation to crime. Found dependencies can then help, for instance, to inform urban design decisions on how to design the urban landscape to improve safety in the city.

At the same time, with the rise of open–source data and increasing collaborations, the same model can be applied to other cities all over the world. Cities and their population differ from country to country, from culture to culture. Therefore it is necessary to understand in what contexts certain theories hold, thus advancing knowledge in the area of urban crime.

Crowdsourcing safety perception towards people: In Chapter 5 I have crowdsourced safety perceptions towards single people by using online images. While results suffered from a number of limitation, I have clearly shown the approaches potential to crowdsource perceptions, as for instance safety. As perceptual 'soft data' has been difficult to grasp on large scale, the method offers a solution. With my results and the increasing significance of crowdsourcing in modern society, I clearly see the method's potential to be refined and re–applied also for other cities.

In terms of refinement, I am very interested to include groups instead of individual people in a next step into the study, as we barely encounter individuals, but groups of people in the city. Matei et al. (2001) suggests that different group sizes and compositions, defined by age, gender and ethnicity for instance, affect our safety perceptions. My presented methodology can be therefore expanded to include groups of people, defined by their size and composition.

Besides the expansion of this study, I want to further investigate the limitations of crowdsourcing as a method. I have detected a lack of crowd–control as a major challenge of online crowdsourcing, biasing results and limiting study outcome. To improve my methodology it is crucial for future work to find ways to minimize these limitations. There exist different approaches to engage and motivate specific crowds, based on gamification for instance (von Ahl and Dabbish, 2008). Still, it is not clear *who* they are attracting. Different strategies, such as gamification, social cause and rewards, need to be explored and analysed what type of crowd each strategy is most successful in engaging, and towards what type of task.

Crowdsourcing safety perception towards urban places: In Chapter 6 I have crowdsourced safety perceptions towards urban places, defined by the built environment and single people inhabiting it. While results suffered from limitations, I show clearly the method's potential to crowdsource safety perception of urban places. As important urban perceptions are in terms of how people use a city, as difficult the data has been to collect quantitatively for researchers and urban planners up till now. Using here presented outcome, the method can now be refined, expanded and re–applied in various ways.

As a next step I am very interested to include groups of people instead of individual people, as it represents the reality in cities more accurately, to see how this affects outcome. At the same time, I will re–apply the study in other cities than London, as architecture and visual appearance of cities all over the world differ according to culture and history, which might have an impact on safety perceptions. Therefore I want to expand the study for several cities that are represented on Google Street View, such as Beijing, New York City and Sao Paolo, and include them in the work.

Furthermore, by including a textbox I have shown that proposed methodology offers opportunities to crowdsource survey–like data, such as personal comments, that go beyond user–votings in a quantitative way, even if it was very generic. To receive more comments on specific questions, I am interested in exploring this approach further in terms of its structure and open the door to natural language (automated) processing (Simm et al., 2010; Nasa et al., 2010) of that data too.

Bibliography

Adams, R. (2008). The englishness of english punk: Sex pistols, subcultures and nostalgia. *Popular Music and Society*, 31(4):469–488.

Adepeju, M., Cheng, T., Shawe-Taylor, J., and Bowers, K. (2015). A new metric of crime hotspots for operational policing. In *Proc of GISRUK*.

AMT (2010). `https://www.mturk.com/mturk/welcome`. Accessed: 2016-04-29.

Anderson, M., MacDonald, J., Bluthenthal, R., and Ashwood, J. (2013). Reducing crime by shaping the built environment with zoning: An empirical study of los angeles. *University of Pennsylvania Law Review*, 161:699–756.

Apple (2015). `http://www.apple.com/uk/watch/`. Accessed: 2016-04-29.

Ashby, D. and Longley, P. (2005). Geocomputation, geodemographics and resource allocation for local policing. *Transactions in GIS*, 9(1):53—-72.

Bannister, J. and Fyfe, N. (2001). Introduction: fear and the city. *Urban Studies*, 38(5-6):807–813.

Beck, U. (1992). *Risk society: towards a new modernity*. London, Sage.

Bendler, J., Brandt, T., Wagner, S., and Neumann, D. (2014). Investigating crime-to- twitter relationships in urban environments—facilitating a virtual neighbourhood watch. In *Proc of European Conference on Information Systems (ECIS)*.

Bettencourt, L., Lobo, J., Helbing, D., Kuehnert, C., and West, G. (2007). Growth, innovation, scaling, and the pace of life in cities. *PNAS - Proceedings of the National Academy of Sciences of the United States of America*, 104(17):7301–7306.

Bogomolov, A., Lepri, B., Staiano, J., Oliver, N., Pianesi, F., and Pentland, A. (2014). Once upon a crime: Towards crime prediction from demographics and mobile data. In *ICMI 2014*.

Bowers, K. and Johnson, S. (2005). Domestic burglary repeats and space-–time clusters: The dimensions of risk. *European Journal of Criminology*, 2(1):67—92.

Boyd, D. and Crawford, K. (2012). Critical questions for big data. *Information, Communication and Society*, 15(5):662–679.

Brantingham, P. J. and Brantingham, P. L. (1981). *Environmental criminology*. SAGE Publications.

Braun, V. and Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(3):77–101.

Brown, M. and Polk, K. (1996). Taking fear of crime seriously: the tasmanian approach to community crime prevention. *Crime and Delinquency*, 42(3):398–420.

Brown, P., Hirschfeld, A., and Batey, P. W. (2000). Adding value to census data: Public sector applications of the super profiles geodemographic typology. *Journal of Cities and Regions*, 10:19–32.

Cambridge (2000). `http://dictionary.cambridge.org/dictionary/english/soft-data`. Accessed: 2016-04-29.

Chainey, S., Reid, S., and Stuart, N. (2002). When is a hotspot a hotspot? a procedure for creating statistically robust hotspot maps of crime. *Innovations in GIS 9 Socio-economic Applications of Geographic Information Science*.

Chainey, S., Tompson, L., and Uhlig, S. (2008). The utility of hotspot mapping for predicting spatial patterns of crime. *Security Journal*, (21 (1–2)):4–28.

Cheng, T. and Adepejue, M. (2013). Detecting emerging space-time crime patterns by prospective stss. In *Proc of International Conference of GeoComputation*.

Christens, B. and Speer, P. (2005). Predicting violent crime using urban and suburban densities. *Behavior and Social Issues*, (14):113–127.

Christin, D., Roßkopf, C., and Hollick, M. (2013). usafe: A privacy-aware and participative mobile application for citizen safety in urban environments. *Pervasive and Mobile Computing*, 9(5):695–707.

Clarke, V., Hayfield, N., and Huxley, C. (2012). Lesbian, gay, bisex- ual and trans appearance and embodiment: A critical review of the psychological literature. *Psychology of Sexualites Review*, 3(1):51–70.

Clickworker (2014). `https://www.clickworker.com`. Accessed: 2016-04-29.

Cohen, J. (1988). Statistical power analysis for the behavioral sciences (2nd ed.). *Current Directions in Psychological Science*, 1(3):98–101.

Cozens, P., Neale, R., Hillier, D., and Whitaker, J. (2004). Tackling crime and fear of crime while waiting at britain's railway station. *Journal of Public Transportation,*, 7(3):23–41.

Crowdflower (2014). `http://www.crowdflower.com`. Accessed: 2016-04-29.

Dasgupta, N. and Greenwald, A. (2001). On the malleability of automatic attitudes: Combating automatic prejudice with images of admired and disliked individuals. *Journal of Personality and Social Psychology*, 81(5):800–814.

Dasgupta, N., Greenwald, A., McGhee, and Banaji, M. (2000). Automatic preference for white americans: Eliminating the familiarity explanation. *Journal of Experimental Social Psychology*, (36):316–328.

Davies, T. and Johnson, S. (2015). Examining the relationship between road structure and burglary risk via quantitative network analysis. *Quantitative Criminology*, (31):481—507.

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

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

Doran, B. and Burgess, M. (2012). *Putting Fear of Crime on the Map: Investigating Perceptions of Crime Using Geographic Information Systems*. Springer.

Eagle, N. and Macy, M. (2010). Network diversity and economic development. *Science*, (1029).

Eck, J., Chainey, S., Cameron, J., Leitner, M., and Wilson, R. (2005). Mapping crime: Understanding hot spots. *Special Report NIJ*.

Ericsson, K. and Simon, H. (1980). Verbal reports as data. *Psychological Review*, 87(3):215—251.

EyeOnMe (2014). `https://itunes.apple.com/gb/app/eyeonme/id718261113?mt=8`. Accessed: 2016-04-29.

Farrall, S., Bannister, J., Ditton, J., and Gilchrist, E. (1997). Questioning the measurement of the 'fear of crime'. *British Journal of Criminology*, 37(4):658–679.

Felson, M. and Clarke, R. (1998). *Opportunity Makes the Thief: Practical theory of crime prevention*. Home Office.

Felson, M. and Poulsen, E. (2003). Simple indicators of crime by time of day. *International Journal of Forecasting*, (19):595–601.

Field, A., Miles, J., and Field, Z. (2012). *Discovering Statistics using R*. SAGE Publications Ltd.

Fitbit (2012). `https://www.fitbit.com/uk`. Accessed: 2016-04-29.

Fiverr (2014). `http://fiverr.com`. Accessed: 2016-04-29.

Flatley, J. (2015). Crime in england and wales: Year ending december 2015. *Office for National Statistics – Statistical bulletin*.

Fountain, M. (2012). *Fear of Crime Survey 2012*. QA research.

Fowler, E. (1987). Street management and city design. *Social Forces*, 66(2):365–389.

Furniss, D., Blandford, A., and Curzon, P. (2011). Confessions from a grounded theory phd: Experiences and lessons learnt. In *Proc of CHI 2011*.

Furstenber, F. (1971). Public reaction to crime in the streets. *The American Scholar*, The fear of crime:3–12.

Gale, C., Singleton, A., and Longley, P. (2015). Profiling burglary in london using geodemographics. In *Proc of GISRUK*.

Gehl, J. (2010). *Cities for People*. Island Press.

Gerber, M. (2014). Predicting crime using twitter and kernel density estimation. *Decision Support Systems*, 61:115—125.

GhettoTracker (2013). `http://www.citylab.com/tech/2013/09/enough-already-avoid-ghetto-apps/6776/`. Accessed: 2016-04-29.

GLA (2011). `http://data.london.gov.uk/census/`. Accessed: 2016-04-29.

Glaeser, E.L., S. B. (1999). Why is there more crime in cities? *Journal of Political Economy*, 107(6):225–258.

Gold, S. (1991). Ethnic boundaries and ethnic entrepreneurship: A photo-elicitation study. *Visual Sociology*, 6(2):9–22.

Google (2016). `https://www.maps.google.com/`. Accessed: 2016-04-29.

Greenwald, A., McGhee, D., and Schwartz, J. (1998). Measuring individual differences in implicit cognition: The implicit association test. *Journal of Personality and Social Psychology*, 74(6):1464–1480.

Hair, J. F. J., Anderson, R. E., Tatham, R. L., and Black, W. C. (1995). *Multivariate Data Analysis (3rd ed)*. New York: Macmillan.

Hall, E. (1966). *The hidden Dimension*. Anchor Books.

Hallman, K. K., Kenworthy, N. J., Diers, J., Swan, N., and Devnarain, B. (2013). The contracting world of girls at puberty: Violence and gender-divergent access to the public sphere among adolescents in south africa. *Poverty, Gender, and Youth Working Paper*, (25).

Hayes, N. (2000). *Doing psychological research: gathering and analyzing data*. Open University Press.

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

Hillier, B. and Hanson, J. (1984). *The Social Logic of Space*. Cambridge University Press.

Hillier, B. and Sahbaz, O. (2009). An evidence based approach to crime and urban design. *Designing Sustainable Cities: Decision-making Tools and Resources for Design*, pages 163–186.

Hirschfeld, A. and Bowers, K. (1997). The effect of social cohesion on levels of recorded crime in disadvantaged areas. *Urban Studies*, 8:1275–1295.

HMIC (2014). `http://www.justiceinspectorates.gov.uk/hmic/wp-content/uploads/crime-recording-making-the-victim-count.pdf`. Accessed: 2016-12-01.

Hollway, W. and Jefferson, T. (1997). The risk society in an age of anxiety: situating fear of crime. *The British Journal of Sociology*, 48(2):255.

Hubert, L. J. and Golledge, R. G. (1982). Measuring association between spatially defined variables: Tjostheim's index and some extensions. *Geographical Analysis*, (14):273–278.

ICA (2000). `http://www.cpted.net`. Accessed: 2016-04-29.

Ipeirotis, P. (2010). Demographics of mechanical turk (working paper).

iWitness (2014). `http://iwitness.usc.edu/SFI/`. Accessed: 2016-04-29.

Jackson, J. (2005). Validating new measures of the fear of crime. *Int. J. Social Research Methodology*, 8(4):297–315.

Jacobs, J. (1961). *The Death and Life of Great American Cities*. Random House Inc.

Jansson, K. (2006). *British Crime Survey: Measuring crime for 25 years*.

Johnson, S. D. and Bowers, K. J. (2004). The burglary as clue to the future: The beginnings of prospective hot–spotting. *European Journal of Criminology*, 1(2):237–255.

Katz, C. M. and Webb, V. J. e. a. (2003). Fear of gangs: a test of alternative theoretical models. *Justice Quarterly*, 20(1):95–130.

Kazai, G., Kamps, J., and Milic-Frayling, N. (2012). The face of quality in crowdsourcing relevance labels: Demographics, personality and labeling accuracy. In *Proc of CIKM'12*.

Kounadi, O., Lampoltshammer, T., E., G., Sitko, I., and Leitner, M. (2015). Exploring twitter to analyze the public's reaction patterns to recently reported homicides in london. *PLoS ONE*, 10(3).

Lam, B., Chan, Y., Whittle, J., Binner, J., Garton, L., and Frankova, K. (2011). Voice your view: An inclusive approach to civic engagement. In *Proc of the 6th International Conference on Inclusive Design (Include 2011)*.

Lane, J. and Meeker, J. W. (2003). Fear of gang the crime: a look at three theoretical models. *Law and Society Review*, 37(2):425–456.

Legendre, P. (1993). Spatial autocorrelation: Trouble or new paradigm? *Ecology*, 74(6):1659–1673.

Lianos, M. and Douglas, M. (2000). Dangerization and the end of deviance: the institutional environment. *British Journal of Criminology*, 40(2):261–278.

LOAC (2011). London output area classification. `https://data.london.gov.uk/dataset/london-area-classification`. Accessed: 2016-09-29.

Longley, P. (2005). Geographical information systems: a renaissance of geodemographics for public service delivery. *Progress in Human Geography*, 29(1):57–63.

Lynch, K. (1960). *The Image of the City*. MIT Press.

MacKerron, G. and Bryson, A. (2013). Are you happy while you work? *Centre for Economic Performance - LSE.*

MacKerron, G. and Mourato, S. (2013). Happiness is greater in natural environments. *Global Environmental Change*, 23(5):992–1000.

Malleson, N. and Andresen, M. (2014). The impact of using social media data in crime rate calculations: shifting hot spots and changing spatial patterns. *Cartography and Geographic Information Science*, 42(2).

Malleson, N. and Andresen, M. (2016). Exploring the impact of ambient population measures on london crime hotspots. *Journal of Criminal Justice*, 46:52—-63.

Mappiness (2013). Mappiness. `http://www.mappiness.org.uk`. Accessed: 2016-09-29.

Marshall, S. (2012). Science, pseudo-science and urban design. *Urban Design Int*, 17(4):257—-271.

Matei, S., Ball-Rokeach, S., and Linchuan Qiu, J. (2001). Fear and misperception of los angeles urban space – a spatial-statistical study of communication-shaped mental maps. *Communication Research*, 28(4):429–463.

Mavros, P., Austwick, M., and Hudson-Smith, A. (2016). Geo–eeg: Towards the use of eeg in the study of urban behaviour. *Applied Spatial Analysis and Policy*, pages 1–22.

Mavros, P., Coyne, R., Roe, J., and Aspinall, P. (2012). Engaging the brain – implications of mobile eeg for spatial representation. In *Proc of eCAADe 30*.

Merry, E. (1981). *Urban danger: life in a neighbourhood of strangers*. Philadelphia, PA, Temple University press.

Milgram, S. (1977). *The individual in a social world: essays and experiments*. Reading, Mass.: Addison-Wesley Pub. Co.

Miller, J. (2008). Impact of situational factors on survey measured fear of crime. *International Journal of Social Research Methodology*, 11(4):307–325.

Milo, M., Richards, S. C., Parang, S., Lonesome, M., and Bhattacharjee, T. (2012). Crime hotspot tracking and geospatial analysis in merseyside, uk. *CS/STAT 5525, Data Analytics*, Fall.

Mirrlees-Black, C. and Allen, J. (1998). *Concern about crime: Findings from the 1998 British Crime Survey. Research Findings No 83*. London, Home Office Research, Development and Statistics Directorate.

Naik, N., Philipoom, J., Raskar, R., and Hidalgo, C. (2014). Streetscore - predicting the perceived safety of one million streetscapes. In *CVPR Workshop on Web-scale Vision and Social Media*.

Nasa, B., Binner, J., Simm, W., Ferrario, M. A., Whittle, J., Lam, B. C., and Zhao, Z. (2010). Voiceyourview: mapping public confidence in policing. In *Proc of Horizon Digital Futures 2010*.

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

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

Neter, J., Wasserman, W., and Kutner, M. H. (1989). *Applied Linear Regression Models*. Homewood, IL: Irwin.

Newman, P. (1972). *Defensible Space: Crime Prevention Through Urban Design*. Macmillian Pub Co.

Nold, C. (2009). Emotional cartography: Technologies of the self. `http://www.emotionalcartogra-phy.net`. Accessed: 2016-04-29.

NYU (2016). Cusp. `http://http://cusp.nyu.edu`. Accessed: 2016-12-11.

Oc, T. and Tiesdell, S. (1997). *Safer city centres: reviving the public realm*. SAGE Publications Ltd.

ONS (2011). Methodology note for the 2011 area classification for output areas. `http://www.ons.gov.uk`. Accessed: 2016-09-29.

ONS (2015). `http://www.ons.gov.uk/ons/index.html`. Accessed: 2016-04-29.

OpenGeodemographics (2016). http://www.opengeodemographics.com. Accessed: 2016-09-29.

Orange (2013). http://www.d4d.orange.com/en/home. Accessed: 2016-04-29.

Pacione, M. (2003). Urban environmental quality and human wellbeing—a social geographical perspective. *Landscape and Urban Planning*, 65(1–2):19–30.

Pain, R. X. (2001). Gender, race, age and fear in the city. *Urban Studies*, 38(5–6):899–913.

Painter, K. (1996). The influence of street lighting improvements on crime, fear and pedestrian street use, after dark. *Landscape and Urban Planning*, 35(2-3):193–201.

Pan, Y. and Jackson, R. T. (2008). Ethnic difference in the relationship between acute inflammation and and serum ferritin in us adult males. *Epidemiology and Infection*, (136):421–431.

PanicButton (2014). https://panicbutton.io. Accessed: 2016-04-29.

Pantazis, C. (2000). 'fear of crime', vulnerability and poverty. *The British Journal of Criminology*, 40(3):414–436.

Park, A. (2008). Modelling the role of fear of crime in pedestrian navigation.

Park, A., H., E., Spicer, V., Cheng, C., Brantingham, P., and Sixsmith, A. (2011). Testing elderly people's fear of crime using a virtual environment. In *Intelligence and Security Informatics Conference (EISIC), 2011 European*, pages 63–69.

Paulos, E. and Goodman, E. (2004). The familiar stranger: Anxiety, comfort, and play in public places. In *Proc of CHI'04*.

Paynich, R. (2013). Identifying high crime areas. *International Association of Crime Analysts*, (2).

Pew (2014). http://www.pewinternet.org/files/2014/04/PIP_Seniors-and-Tech-Use_040314.pdf. Accessed: 2016-04-29.

PoliceUK (2013a). `http://data.police.uk`. Accessed: 2016-04-29.

PoliceUK (2013b). `https://data.police.uk/about/#location-anonymisation`. Accessed: 2016-04-29.

Prasetyo, P., Gao, M., Lim, E., and Napa Scollon, C. (2013). Social sensing for urban crisis management: The case of singapore haze. In *Proc of SocInfo 2013*, pages 478–491.

Quattrone, G., Capra, L., and De Meo, P. (2015). There‘s no such thing as the perfect map: Quantifying bias in spatial crowd-sourcing datasets. In *Proc of CSCW'15*.

Quercia, D., O'Hare, N., and Cramer, H. (2014). Aesthetic capital: What makes london look beautiful, quiet, and happy. In *CSCW'14*.

Quercia, D., Pesce, J., and Almeida, V. (2013). Psychological maps 2.0: A web engagement enterprise starting in london. In *Proc of WWW'13*.

Quercia, D., Schifanella, R., Aiello, L. M., and McLean, K. (2015). Smelly maps: The digital life of urban smellscapes. In *Proc of ICWSM'15*.

Ray, J. C. (1971). *Crime Prevention Through Environmental Design*. SAGE Publications.

RedPanicButton (2014). `http://www.redpanicbutton.com`. Accessed: 2016-04-29.

Rogerson, P. A. (2001). *Statistical methods for geography*. London: Sage.

Rose, G. (2007). *Visual Methodologies: An Introduction to the Interpretation of Visual Materials*. SAGE Publications.

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

Ross, J., Irani, I., Silberman, M., Zaldivar, A., and Tomlinson, B. (2010). Who are the Crowdworkers? Shifting Demographics in Mechanical Turk. In *Proceedings of CHI'10*.

Sahbaz, O. and Hiller, B. (2007). The story of the crime: functional, temporal and spatial tendencies in street robbery. In *Proc of 6th International Space Syntax Symposium, Istanbul*, pages 04–14.

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

Saroglou, V. (2014). *Religion, Personality, and Social Behavior*. Routledge – Psychology Press.

Schmidt, C. (1977). Influence of land use diversity upon neighborhood success: An analysis of jacobs' theory. *The Annals of Regional Science*, 11(1):53–65.

SchoolGuard (2014). `http://www.guard911.com/schoolguard/`. Accessed: 2016-04-29.

Sen, S., Macademia Team: Giesel, M., Gold, R., Hillmann, B., Lesicko, M., Naden, S., Russell, J., Wang, Z., and Hecht, B. (2015). Turkers, scholars, "arafat" and "peace": Cultural communities and algorithmic gold standards. In *Proc of CSCW'15*.

Shannon, C. E. and Weaver, W. (1949). *The Mathematical Theory of Communication*. University of Illinois Press, Urban.

Simm, W., Ferrario, M., Piao, S., Whittle, J., and Rayson, P. (2010). Classification of short text comments by sentiment and actionability for voiceyourview. In *Proc of SocialCom 2010*.

Singleton, A. and Longley, P. (2015). The internal structure of greater london: a comparison of national and regional geodemographic models. *Geography and Environment*, 2(1):69–87.

SINTEF (2013). `www.sciencedaily.com/releases/2013/05/130522085217.htm`. Accessed: 2016-04-29.

SketchFactor (2014). `http://www.sketchfactor.com`. Accessed: 2016-04-29.

Skogan, W. (1986). Fear of crime and neighborhood change. *Crime and Justice*, 8:203–229.

Skogan, W. (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.

Smith Clarke, C., Mashhadi, A., and Capra, L. (2014). Poverty on the cheap: estimating poverty maps using aggregated mobile communication. In *Proc of CHI'14*, pages 511–520.

Snyder, M. (2007). The impact of recent immigration on the london economy. Technical report, London School of Economics and Political Science.

Solymosi, R., Bowers, K., and Fujiyama, T. (2015). Mapping fear of crime as a context-dependent everyday experience that varies in space and time. *Legal and Criminological Psychology*, 20:193–211.

Song, W. and Daqian, L. (2013). Exploring spatial patterns of property crime risks in changchun, china. *International Journal of Applied Geospatial Research*, 4(3):80–100.

Tan, S. and Haining, R. (2009). An urban study of crime and health using an exploratory spatial data analysis approach. In *Proc of ICCSA'09*, volume 1, pages 269–284.

Taxiaviso (2014). `https://taxiaviso.com/home/`. Accessed: 2016-04-29.

Taylor, R. B. and Hale, M. (1986). Testing alternative models of fear of crime. *The Journal of Criminal Law and Criminology*, 77(1):151–189.

TelecomItalia (2012). `http://www.telecomitalia.com/tit/en/bigdatachallenge.html`. Accessed: 2016-04-29.

Telefonica (2014). `http://dynamicinsights.telefonica.com/tag/datathon-for-social-good/`. Accessed: 2016-04-29.

Tjostheim, D. (1978). A measure of association for spatial variables. *Biometrika*, (65,1):109–114.

Tobler, W. (1970). A computer movie simulating urban growth in the detroit region. *Economic Geography*, 46:234–240.

Tuan, Y. (2001). Space and place: The perspective of experience.

Tukey, J. (1949). Comparing individual means in the analysis of variance. *Biometrics*, 5(2):99–114.

Tulloch, M. (2000). The meaning of age differences in the fear of crime. *The British Journal of Criminology*, 40(3):451–467.

UCL (2016). Uclethics. `http://ethics.grad.ucl.ac.uk`. Accessed: 2016-12-11.

Uhlmann, E., Dasgupta, N., Elgueta, A., Greenwald, A., and Swanson, J. (2002). Subgroup prejudice based on skin color among hispanics in the united states and latin america. *Social Cognition*, 20(3):198–225.

UNDESA (2014). Definition of youth.

Vernon, R., Sutherland, C., Young, A., and Hartley, T. (2014). Modeling first impressions from highly variable facial images. *PNAS*, 111(32):3353–3361.

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

Wakamiya, S., Lee, R., and Kazutoshi, S. (2013). Social-urban neighborhood search based on crowd footprints network. In *Proc of SocInfo 2013*, pages 429–472.

Walkonomics (2014). `http://www.walkonomics.com`. Accessed: 2016-04-29.

Wang, D., Ding, W., Lo, H., Stepinski, T., Salazar, J., and Morabito, M. (2013). Crime hotspot mapping using the crime related factors—a spatial data mining approach. *Applied Intelligence*, 39(4):772—-781.

Wang, X., Gerber, M., and Brown, D. E. (2012). Automatic crime prediction using events extracted from twitter posts. *Social Computing, Behavioral - Cultural Modeling and Prediction*, 7227(4):231–238.

Warr, M. (1984). Fear of victimization: why are women and the elderly more afraid. *Social Science Quarterly*, 65(3):681–702.

Warr, M. and Ellison, C. (2000). Rethinking social reactions to crime: personal and altruistic fear in family households. *American Journal of Sociology*, 106(3):551–578.

Weicher, J. (1973). A test of jane jacobs' theory of successful neighborhoods. *Journal of Regional Science*, 13(1):29–40.

Whittle, J., Simm, W., and Ferrario, M. (2010). Voiceyourview: collecting real-time feedback on the design of public spaces. In *Proc of UbiComp 2010*.

WHO, U. H. (2010). *Hidden cities: unmasking and overcoming health inequities in urban settings*. WHO, Library Cataloguing-in-Publication Data.

Williams, M., Burnap, P., and Sloan, L. (2016). Crime sensing with big data: The affordances and limitations of using open–source communications to estimate crime patterns. *British Journal of Criminology*.

Willis, J. and Todorov, A. (2006). First impressions: Making up your mind after 100 ms exposure to a face. *Psychological Science*, 17(7):592—-598.

Wilson, J, Q. and Kelling, G. L. (1982). "broken windows: The police and neighborhood safety". *Atlantic Monthly*, 249(3):29–38.

Wobbrock, J., Findlater, L., Gergle, D., and Higgins, J. (2011). The aligned rank transform for nonparametric factorial analyses using only anova procedures. In *Proc of CHI'11*.

Wolfe, M. and Mennis, J. (2012). Does vegetation encourage or suppress urban crime? Evidence from Philadelphia, PA. *Landscape and Urban Planning*, 108:112–122.

Wood, E. (1967). *Social aspects of housing and urban development*. Department of Economic and Social Affairs, United Nations, New York.

Zako, R. (2009). Young people's gatherings. *Urban Public Realm*, (66).