

Analysis and mapping of crime perception: A quantitative approach of sketch maps.

Mariana Vallejo Velazquez ¹, Ourania Kounadi ¹[0000-0002-5998-7343], and Andrea Podor ²

¹ Faculty of Geo-Information Science and Earth Observation (ITC), University of Twente,
Enschede, The Netherlands

² Institute of Geoinformatics, Óbuda University Alba Regia Technical Faculty, Székesfehérvár
8000, Hungary,
mvv_geo@hotmail.com
o.kounadi@utwente.nl
podor.andrea@amk.uni-obuda.hu

Abstract. Evidence exists that people's perception of crime is not often consistent with the actual incidents statistics, and there is a tendency of underestimating or overestimating safety. We examine a phenomenon called the *crime perception gap* via participatory geographical information derived from sketch maps. The study area is Budapest, Hungary for which data were collected via a participatory platform in 2017 on the perception of safe and unsafe places. The methodology consisted of three stages; exploratory modeling, the spatial delineation of the gap, and the spatial exploration of inaccurate perceptions in relation to their surrounding environment. In stage one, we found that the variable with the highest impact on perception is the daily route. The further away a place is to personal routes the more likely it is to be perceived as unsafe. In the second stage, we computed and mapped the perceptual accuracy. The overall perceptual accuracy was as low as 39%, while many safe places were wrongly perceived as unsafe (also unsafe ones as safe). In the third stage, we identified that significant spatial patterns seem to have a *diffusion effect* on people's perception. For example, a safe place could be perceived as unsafe because the neighboring places are crime hotspots (and vice versa). We argue that misperception of crime can have repercussions on peoples' lifestyles, affect social behavior and spatial and economic dynamics. Thus, spatial analysis and mapping can be used to support police agencies in the development of strategies to reduce the misperception of crime.

Keywords: Crime Perception Gap, Spatial Statistics, Spatial Modelling, Cognitive Mapping, Sketch Maps, Participatory GIS (PGIS).

1 Introduction

Cognitive mapping is the process of developing a mental map, based on the collection of information by sensorial perception. The geometry and attributes of each individual

cognitive map are shaped by internal and external factors. The main tangible representation of a cognitive map is the sketch map as the main mapping method to graphically depict the spatial knowledge of individuals. Relative location, geometric attributes, as well as impressions and beliefs about places can be portrayed in these representations.

Sketch mapping is a recurrent method for the collection of data about people's perception. Researches used it to study diverse social geographical matters such as emergency management, hazard planning, land-use planning, and community safety [1].

Crime studies is one of the research areas in which structured sketch maps are utilized with the aim of collecting data about fear of crime [2,3] and perception of crime [3,4]. Different internal and external factors shape the perception of crime. Crime perception analysis has been approached in two ways: with cross-sectional and longitudinal statistical studies. While the first one analyses a single sample of data, the second compares different samples from different given times, which allows tracing the changes over time. These studies explored various topics such as urban areas [5], business areas [6], neighborhoods [7], university campus [8] and daily routine areas [9].

Although there are several theories that explain the factors that sway the perception of crime, the spatial component of it has not been explored in depth. Spatial characterization of those factors could result in a better understanding of the perception of unsafe areas. The relevance of this lies in the fact that perception is not always similar to reality: perceptions of crime often mismatch the actual crime statistics. This disparity is known as the crime perception gap and it arises when a person has an inaccurate insight of safety. In specific, the crime perception gap is defined as the difference between the level of unsafety that is conceived by a person and the actual level of unsafety based on actual crime incidents [10].

Regarding unsafe areas perceived as safe, this is usually related to heuristics or incivilities that some people make in areas where they are less familiar with or where they are not related. The fact that unsafe areas perceived as safe happens usually in people's own neighborhood, can be explained by the "endowment effect" which consists in assigning a higher or better value to the objects we possess, than to the same objects that we do not own [11]. In crime perception, this can be applied when people tend to characterize their neighborhood as safe, under the assumption that better conditions exist in areas where people live than other unknown areas. People tend to have a perceptual bias due to a feeling of attachment toward their own community or neighborhood [12].

These social misperceptions can have an impact on different aspects. For example, if people are not aware of the high risk of victimization, thus appropriate precautions are not being taken and the probability of an attack might increase. Regarding the inaccurate perception of a safe place as being unsafe, the impact can be on a bigger scale and its effects can last for a longer time, as it has repercussions on people's lifestyle [13], health due to anxiety [14], social behavior and the spatial and economic dynamics [15]. A high crime perception can restrict individual activity spaces due to the avoidance of unsafe areas or streets at certain hours. On these areas labeled perceptually as unsafe, fewer people would transit, eventually forcing the relocation of shops, restaurants, or businesses that could be affected, which could lead to a reorganization of the spatial activities.

Thus, there is a need to increase perception accuracy with localized strategies that narrow the gap. Although police agencies have developed actions to address this issue, they have mainly focused on reducing the fear of crime [16]. They are particularly focusing on the inaccurate perception of high crime, in which people believe that the level of crime incidents is high, whereas in reality, it is low. But then there is still the need to narrow the gap of an inaccurate perception of low crime in existing crime hotspots wherein the people are not aware of the risk of victimization.

The term “accuracy/inaccuracy of crime perception” is used in this paper as the state of consistency between what is perceived and the reality defined by objective measurements. As perception cannot be described as “right” or “wrong”, the concept of accuracy is employed to establish whether or not the perceived attribute matches the measured value.

The aim of this study is to quantitatively examine structured sketch maps and to analyze and map crime perception. In order to fulfill this, the specific objectives (SOs) are:

1. *To analyze the location of perceived unsafe areas in relation to a) the distribution of crime incidents and b) people’s activity spaces.*
2. *To determine the accuracy of people’s crime perception and to map its spatial distribution.*
3. *To explore inaccurate perceptions in relation to their surrounding environments.*

SO1 tests whether personal activity spaces lead to bias in the perception of crime prevalence. SO2 tests (and quantifies) the existence of crime perception gap, as it has been identified in other studies, but from a spatial perspective. SO3 tests whether people’s perception follows the Tobler’s first law of geography [17] (i.e. “*everything is related to everything else, but near things are more related than distant things.*”) and thus transfer perceptions of one place to nearby locations.

The next section (2) contains the description of the datasets, the pre-processing and geo-processing procedures while in section 3, we tackle the first specific objective with exploratory modeling on crime perception. Section 4 addresses the second objective, by quantifying the spatial delineation of perceptual accuracy on crime. Section 5 addresses the third objective with the use of spatial statistics on the places of inaccurate perceptions. In the last two sections (6 and 7) we discuss results and conclude with recommendations for future work.

2 Data: description, pre-processing and geo-processing

The data used in this study is derived from an ongoing participatory online survey¹ conducted at a national level in Hungary by the Institute of Geoinformatics from the Óbuda University. The initial participants were students from the University, afterward, the survey spread out by a snowball effect. The data used in this research was collected

¹ Measuring the fear of crime: <http://bunmegelozes.amk.uni-obuda.hu/MainPageEng.php?ln=1>

in 2017 and is constrained to the city of Budapest². Budapest, is selected as the most crime affected city in Hungary, but not researched extensively [18]. The survey requested to draw a digital structured sketch map over a web-based map, in which the participant indicated the areas that (s)he perceives as unsafe or safe. Also, participants marked with lines their daily routes and were asked to give further information such as their age, sex, postal code where they live, and the main mean of transportation they use.

Table 1 is a summary of the structured sketch maps of the 113 participants. In total, there were 113 participants (39 women, and 74 men), between 18 and 76 years old who drew their daily route(s) and at least one polygon. From the resultant digital sketch maps, three vector files were extracted: perceived safe areas (97 polygons), daily routes (214 lines), and perceived unsafe areas (231 polygons).

Table 1. Summary of the structured sketch maps by gender and sketched element.

Gender	Total participants	Daily routes	Polygons	
			Safe	Unsafe
Women	39	69	31	84
Men	74	128	66	147

Moreover, the Óbuda University provided a CSV file with 60,784 addresses of the recorded crime incidents in Budapest during 2017 (obtained from official records of the police authorities). The addresses were geocoded with the web service Nominatim, a search engine for OpenStreetMap data³. From the data cleaning process, 1,218 records (2%) were deleted due to the lack of an address. Geocoding was run with a set of 59,566 records, from which 58,379 addresses were geocoded, which equals a hit rate of 98%. According to Ratcliffe [19], a minimum geocoding hit rate of 85% is needed to produce an accurate map which reflects the actual distribution of the criminal events.

The original dataset contained crime incidents of spatial or non-spatially-explicit nature such as fraud, crimes against a computer system and data, health-related, misuse of documents and blackmail. Thus, the crime data passed through another filtered process as the research focuses on street crimes, which are the criminal offenses that happen in public places. The data was reduced to 42,805 events.

The geocoded points were spatially joined to the city blocks by processing road data from OpenStreetMap⁴. The point aggregation in blocks was done due to the quality of the geocoding results. For some addresses, the points were located in the centroid of a block, mainly when the address corresponded to a specific public place such as a mall, park, and airport or train station. The additional difficulty is that this type of place tends to accumulate multiple crime incidents. So, in the same pair of coordinates, there could be more than one hundred points. Thus, grouping the points by block allows a characterization of the block in which the place is contained and not of a single point location.

² *Perceptual data* can be shared with potential collaborators/researchers. Please contact the authors if you would like to further/additionally explore the data.

³ Nominatim: <https://nominatim.openstreetmap.org/>

⁴ Open Street Map (OSM): <https://www.openstreetmap.org>

Regarding the first objective on spatial modeling, sketch polygons were split into cells with a rectangular grid in which the centroid of each cell was obtained (Figure 1.A) so that each centroid could represent one data sample.

The approach of the polygon segmentation is suitable for the analysis of sketch maps in the context of perception. As a sketch map is the external representation of an individual cognitive map, it has to be considered that each mental map has a different scale. From the sketched polygons, we observed that the participants drew polygons of various sizes. Some of them drew polygons following the city blocks of the base maps. Other participants drew polygons that do not follow the geometry of the city's administrative boundaries and have a comparably bigger size. To dispel this differentiation and elude generalizations, it was convenient to work with the smallest possible analysis unit. The purpose of segmenting the polygons is to characterize as precise as possible the sketch maps, since, the polygons drawn, are mainly irregular figures that cannot merely be generalized.

The cell's length is 45x45m and was selected based on the smallest drawn polygon. So instead of analyzing 328 polygons, 68,032 cells' centroids that were within the polygons (Figure 1.B) were explored. The centroids' data set includes an identification number, the participant's and polygon's ID, and the type of polygon to which that centroid belongs to, either a safe or unsafe perceived area.

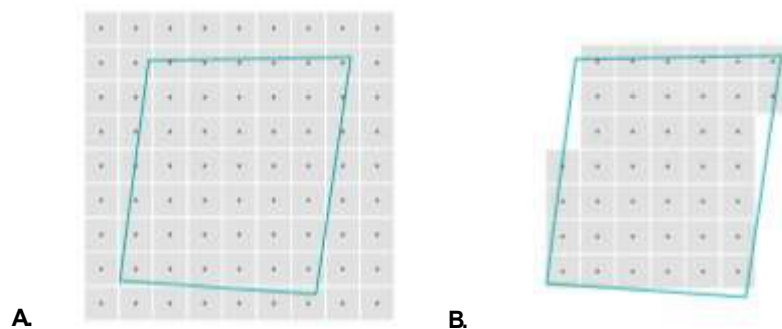


Fig. 1. A) Segmentation of the sketch polygons with a rectangular grid and its centroid. B) Selection of the centroids within the polygons.

For the second and third objective (i.e. the spatial delineation of the perception accuracy and the inaccurate perception in relation to the surrounding environment), we compare the “reference classification” (derived from crime records) and the “perceived classification” (derived from sketch maps) of safe and unsafe areas in Budapest. To perform the comparison both datasets have to be in the same spatial unit. As the reference classification is defined by the actual number of crime events and these were aggregated by blocks, the perceived classification, defined by the sketch polygons, was transformed also in blocks.

3 Modeling of Crime Perception

3.1 Methodology

The first objective is to explore the relationship between the location of the a) people's activities, b) the crime hotspots, and c) the location of the perceived unsafe and safe areas. To address this, we employed a supervised method in which the target (Y) variable is the centroid of the cell with the binary class label of "safe" or "unsafe". The input data (X) were constructed from the spatial analysis between the target and the additional variables derived from people's activities and the locations of crimes. In specific, we used the logistic regression method because the output is not only a resultant class (i.e. the estimated perception of safety) but also an expression of the relationship between the independent variable(s) (X) and the output class (Y). Hence this method is suitable to tackle the first objective of the research, as the problem deals with binary classification (safe and unsafe) and the coefficients of the regression indicate the relationship of the explanatory variables and the dependent variable.

As explained previously labeled polygons of safety perception were disaggregated into cell centroids. Each cell represents a data sample and has a label as being safe or unsafe (Y). From each cell, distance-based measurements were computed to engineer five explanatory variables that consist of the input data (X): a) the participant's neighborhood (postal code area), b) her daily route, c) a crime hotspot, d) a crime spatial outlier, and e) high crime intensity areas. An explanation of the choice of each variable is given in the list below.

1. Neighborhood

The purpose of this variable is to explore whether people tend to perceive their own neighborhood and the surrounding area as safe or unsafe.

2. Daily route

This variable describes if people follow "safe routes" traced in their cognitive maps to avoid high crime perceived areas.

3. Cluster Hotspot & 4. Outlier Hotspot

A hotspot or spatial outlier is a statistically defined location in which local structure is sufficiently unusual. The work of crime analysts and police authorities is predominantly focused on such places and thus we want to see if participants (representing the normal population) are aware of these places and perceive/label them correctly (e.g., a hotspot is an unsafe location). Thus, the local Moran's I was selected to perform hotspot analysis and derive types of spatial local association [20]. The types that are used here are the high crime areas surrounded by high crime areas (cluster hotspot) and the high crime areas surrounded by low crime areas (outlier hotspot).

5. High crime intensity area (HCIA)

We defined as HCIA four blocks with an unusually high number of crime events (between 304 and 614). Such places, and their illegal activities, could be known to the

general public and affect their perception. However, the reality is that if a person transverses areas of high crime density he/she would not necessarily have a higher risk of victimization because the latter is linked to the population density as well.

3.2 Spatial Modeling

The 68,032 samples (cells' centroids within the sketched polygons) of the dataset were divided between training data with a percentage of 80% and testing data with a percentage of 20%. The personal attributes of the participants (age, sex and the main means of transportation) were initially explored in the binary regression analysis. The results showed that the participants' means of transportation was not significant for the 95% selected confidence level, as the p-value was higher than 0.05. Meanwhile, the age and sex explained only 2% of the likelihood's variability of perceiving an unsafe area. Thus these variables were not considered for the final model, as they had not a significant impact on crime perception. Table 2 shows the results of logistic regression.

Table 2. Result of the binary logistic regression (target is the perception as safe/unsafe).

<i>Covariate</i>	<i>Coefficients</i>	<i>e^b</i>	<i>Standard error</i>	<i>z</i>	<i>P> z </i>
<i>Distance to Neighborhood</i>	0.0002	1.0002	8.28 e-05	29.284	0.000
<i>Daily route</i>	0.0010	1.0010	1.93 e-05	53.569	0.000
<i>Cluster hotspot</i>	0.0007	1.0007	1.93 e-05	35.643	0.000
<i>Outlier hotspot</i>	0.0009	1.0009	1.79e-05	52.384	0.000
<i>High crime intensity</i>	0.0001	1.0001	6.55 e-06	12.625	0.000
<i>Constant</i> -3.1573 <i>Pseudo R²</i> 0.312 <i>Log-Likelihood</i> -25,792 <i>Log-Null</i> -37,466					

The coefficients represent the estimated change in the logarithm of the odds of $Y=1$ occurring when all other independent variables are held as constant. For this model $Y=1$ means classifying an area as unsafe, therefore the coefficients are interpreted over this variable. The p-values indicate that the five variables are related to the classification of unsafe areas. The resulting coefficients are explained in terms of their odds ratio which is usually expressed by the exponent of b , e^b (Table 2). When the odds ratio is greater than 1 it means that the odds of getting $Y=1$ increases when the X increases while with values less than 1 the odds of getting $Y=1$ decreases when X decreases. Due to the fact that the odds ratio is not a linear function of the coefficients, it is necessary to estimate the coefficients with the specific number of units X and then get the exponential of the coefficient. In this case, the covariates X were estimated in meters (distances measured) and, therefore, the likelihood is given in reference to one-meter distance.

In order to make the interpretation of the resultant coefficients more meaningful, the units of the independent variable (i.e. the distance in meters) were transformed into the number of blocks. Thus, the likelihood of perceiving an unsafe area relates to the number of blocks away from the target locations. Considering the mean block size in Budapest, it can be estimated the average length of a block as 230 meters. Figure 2 shows how the likelihood of perceiving an area as unsafe changes while moving away from the people's neighborhood, daily route, a crime hot spot, and high crime intensity areas in approximated block size units.

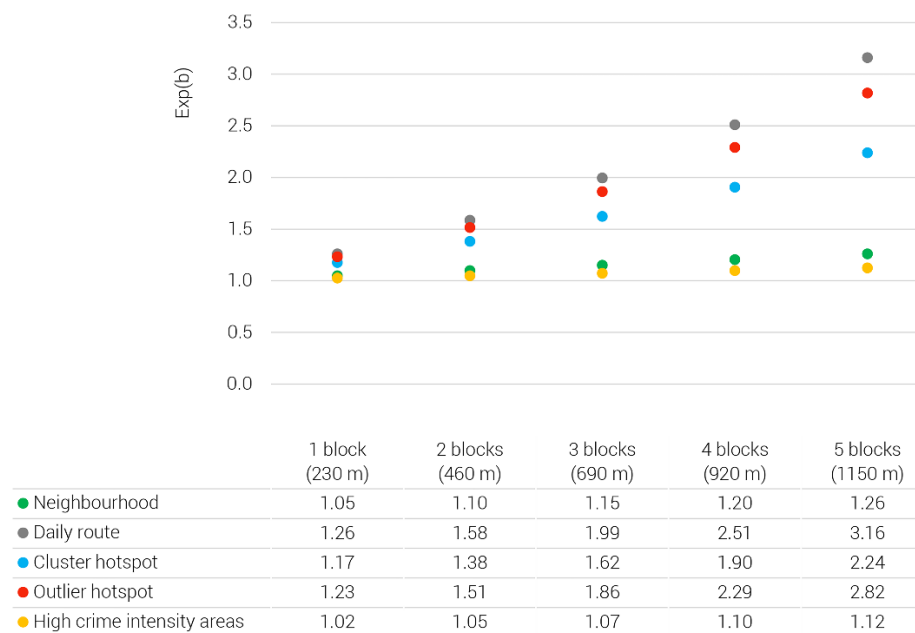


Fig. 2. Exponents of the coefficients e^b obtained for different ranges of distances or number of blocks from the target locations.

The coefficients show that the likelihood of perceiving an unsafe area increases when moving away from the five selected referenced locations. As shown in Figure 2, the increment of the likelihood value is not linear. For the covariates daily route, cluster hotspot, and outlier hotspot, the gradient changes faster than for neighborhood and high crime intensity areas.

The likelihood of perceiving an unsafe area highly increases while moving away from the people's daily route. Participants identified unsafe areas further away from their daily routes, 50% of the participants sketched safe areas in the distance no longer than 200 m. Meanwhile, the unsafe areas were identified in the distance up to 1.2 km by 50% of the participants.

Also, the likelihood increases with increasing distances to peoples' neighborhoods. 50% of the centroids within an identified unsafe areas were less than 1 km away from the participants' neighborhood, meanwhile, half of the centroids of the safe areas were less than 400 m away. In general, the participants identified safe areas closer to their neighborhood.

In the case of the high crime intensity areas, the variation of the likelihood over distance presents a smooth increase. The increment of the likelihood is influenced by the fact that people identified unsafe areas 14 km away from the zone of HCIA that are

located around the city center. This means that people perceived a higher percentage of safe areas around the HCIAs. This could be explained by the high percentage of participants that live in the surrounding areas where the HCIAs are located.

For both types of hotspots, the likelihood increases more or less in the same proportion. For 122 (52.8%) of the sketched unsafe areas at least 50% of their area was within a crime hotspot, and for 40 (17.3%) of them, the entire polygon was contained by a hotspot block. For 62 (63.9%) of the sketched safe areas, at least 50% of their area overlapped with crime hotspots and 10 (10.3%) were entirely within a hotspot.

While the effect of previous variables is either justified by theory (i.e. biased perception on activity spaces) or the data distributions (i.e. HCIA were found only in the center), the effect of hotspots is somehow unexpected (i.e. a negative coefficient sign would have been meaningful). There is a mismatch between reality and perception, which is further analyzed and presented in the next section.

4 Spatial Delineation of Perceptual Accuracy

4.1 Methodology

We delineate the spatial distribution of the level of crime perception accuracy by comparing the class (safe/unsafe) to which each block belongs according to the perceived classification and the reference classification. The perceived classification is defined by counting the sketch maps by type (safe/unsafe) that overlap within one block. Meanwhile, the reference classification of the blocks was defined by crime hotspots. Thus, if a block was labeled as a hotspot then it belongs to the unsafe class, and if it was not (coldspot or insignificant), then it was categorized as safe.

The crime perception gap is represented and classified into the following four types: a) accurate perception of safe area (AS), b) inaccurate perception of safe area (IS), c) accurate perception of unsafe area (AU), and d) inaccurate perception of unsafe area (IU). Figure 3 shows the classification types.

		REFERENCE	
		SAFE	UNSAFE
PERCEPTION	SAFE	Accurate perception of safe area (AS)	Inaccurate perception of safe area (IS)
	UNSAFE	Inaccurate perception of unsafe area (IU)	Accurate perception of unsafe area (AU)

Fig. 3. Types of crime perception accuracy based on safety attributions.

The first step was to count per type (safe/unsafe) the number of participants who sketched a polygon that has at least one cell's centroid within a block. Then, the percentage of participants who classified the block as unsafe from the total number of participants who sketched on that block was calculated.

Obviously, the result ranged from 0 to 100, where 100 indicates that all participants agreed on classifying the block as unsafe, for instance, and zero indicates that everybody agreed on categorizing the block as safe. 50 indicates that the same number of persons identified the block as safe or as unsafe. Thus, when the percentage was higher than 50 the block was labeled as “perceived unsafe”, when it was smaller than 50 it was labeled as “perceived safe”, and when the percentage was 50 the block was “undefined”. Table 3 is an extraction of the blocks’ attribute table to exemplify the way blocks were classified based on the participants’ perception.

Table 3. Example of the block dataset including the “perceived classification”.

Block ID	Hotspot	Participants who classified the block by type		Total	% Unsafe	Perceived Class
		Safe	Unsafe			
1	no	13	6	19	31.5	safe
2	yes	15	36	51	70.6	unsafe
3	yes	3	3	6	50	undefined

Figure 4 shows a scatter plot of the percentage of participants per block who identified it as unsafe and its corresponding number of crime incidents; each point represents a block. This plot shows the presence of a crime perception gap in the study area; as three of the four blocks with the highest amount of crime incidents were identified as safe areas by the majority of participants who sketched over those blocks. Contrary, some blocks were classified as unsafe where there were no reported incidents. On the other hand, there are also blocks that people are aware of the high and low crime rate. Thus, the blocks vector file contains, among other attributes, the values of the perceived and the reference classification. Both values were compared and the blocks were classified into one of the four types of crime perception accuracy (i.e. AS, AU, IS, IU).

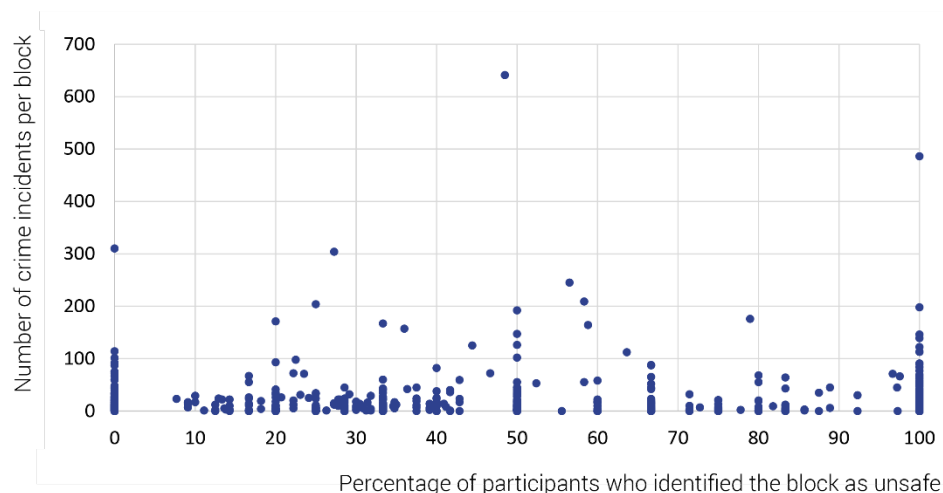


Fig. 4. Crime incidents per block and percentage of participants who identified the block as unsafe.

The next step consisted of defining the level of accuracy or inaccuracy of people's perception. If the block was accurately classified (reference classification = perceived classification), the level of accuracy was defined by the percentage of participants who correctly classified the block by the total number of participants who classified the block. If the block was inaccurately classified (reference classification \neq perceived classification), the inaccuracy level was defined by the percentage of participants who incorrectly classified the block by the total number of participants who classified the block. Based on the percentage values, an ordinal classification was defined to determine three levels: low ($>50\%$ - 65%), medium ($>65\%$ - 85%) and high ($>85\%$ - 100%). For the two accurate classes (AS and AU) this scale represents the proportion of participants who are aware of the safety situation. In this case the blocks that were labeled as low accuracy means that the proportion between the people who were accurate is slightly higher than those who were not. For the case of the two inaccurate classes (IS and IU) the scale represents the proportion of people who are not aware of the crime situation. In this case, the blocks that were classified as high require more attention than those that were labeled as low. Table 4 shows an example of the accuracy type classification and the level of accuracy or inaccuracy.

Table 4. Example of an accuracy type and level classification of 4 blocks.

Block Id	Number of Participants who Classified the Block by Type (%)		Accuracy Type	Level of Accuracy or inaccuracy
	Safe	Unsafe		
1	4 (36.4)	7 (63.6)	AU	Low
2	7 (87.5)	1 (12.5)	AS	High
3	13 (76.5)	4 (23.5)	IS	Medium
4	3 (13.6)	19 (86.4)	IU	High

According to the final classification of blocks, the crime perception gap is identified where the blocks were classified as “inaccurate perception of safe areas” and “inaccurate perception of unsafe areas”, as in these blocks the perception does not correspond with reality. The relevance of distinguishing between these types of inaccuracy lays in the fact that the strategies needed to narrow the perception gap are different for each type of inaccuracy: whereas in the IS people need to be aware of the risk of victimization, in the IU the strategies must be focused on reassuring the people. In order to develop plans of action, it is required to explore the possible causes that explain the inaccuracy in those specific locations.

4.2 Spatial delineation

In total there are 9,655 blocks in Budapest; from which 1,706 lie within the sketched polygons, and thus they were classified by the participants as unsafe or safe. Only these blocks were examined in the crime perception gap analysis. As such, the produced maps depict the part of the city where sketched polygons sprawl (center, east, and south; in about three-quarters of the entire city) and thus leaving out some northern and western

parts. From these classified blocks, 302 are actual crime hotspots and they were classified as “reference” unsafe areas. The rest (1,404), for the purpose of this research, were considered as “reference” safe areas as they are no hotspots.

Figure 5 shows the blocks that were accurately perceived. 37.7% (114) of the blocks identified as hotspots were accurately perceived as unsafe, meanwhile, 37.3% (524) of the non-hotspots were accurately perceived as safe. The map shows some visible clusters of safe and unsafe areas where people are aware of the crime rate. The lightest green areas are those blocks where prevention actions must be taken, as in comparison with the total number of participants who sketched over those blocks, the percentage of those who are aware that the area is a crime hotspot is low.

Figure 6 depicts the blocks that were inaccurately classified, thus this map shows the actual crime perception gap. 54% (163) of the hotspots blocks were inaccurately perceived as safe, and 58.5% (822) of the safe blocks were inaccurately perceived as unsafe blocks. In the center of the city, people tend to have an IS perception, meanwhile, the IU perception happens in the south and southeast part of the city. The map also shows hotspot blocks that were not classified by the participants. These are considered as another block type due to the fact that they are conceptually part of the perception gap. But as they do not have a value in the “perceived classification” attribute and the level of inaccuracy cannot be measured.

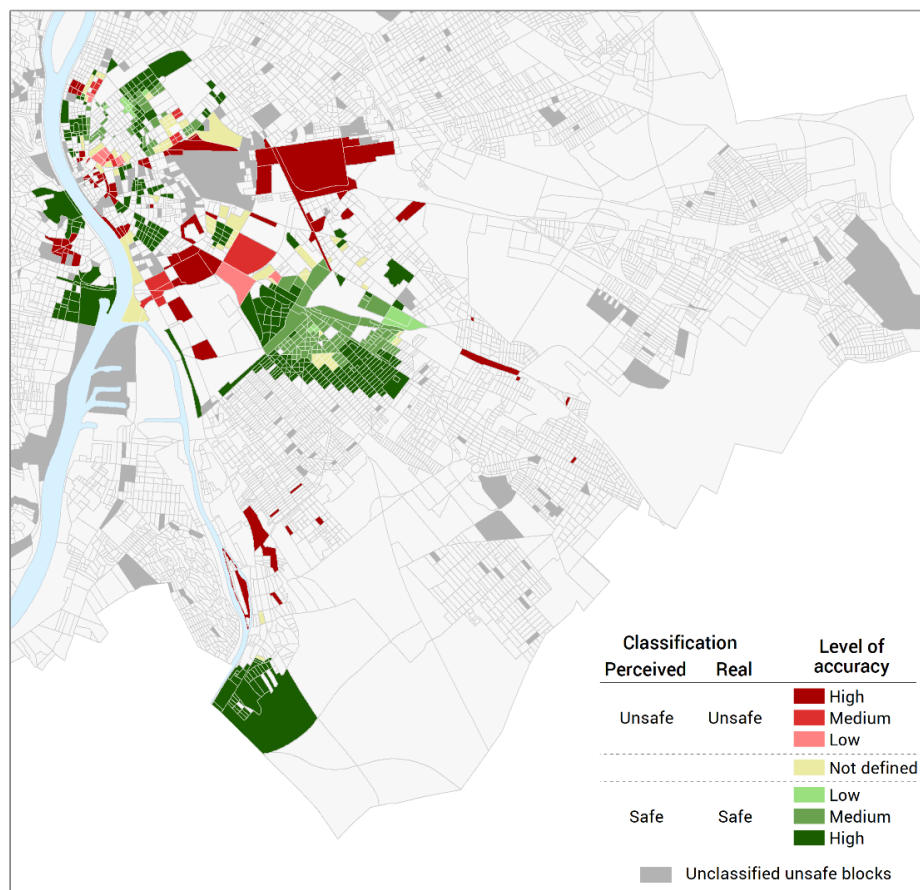


Fig. 5. Level of accurate perception of unsafe (AU) and safe (AS) areas per block.

The accuracy of the participants' perception is presented in Table 5 as a confusion matrix showing the blocks that were correctly and incorrectly classified, as well as the commission and omission errors. Out of the 1,706 classified blocks, 83 were 'not de- fined' due to half of the participants classified those blocks as safe and the other half as unsafe. From the labeled blocks, 61% of the safe blocks were identified as unsafe and 59% of the unsafe blocks were identified as safe. The overall accuracy of the classification is 39%, which is the percentage of accurately classified blocks.

Table 5. Confusion matrix of the crime perception classification.

		Reference			Error of commission
		Safe	Unsafe	Total	
Perception	Safe	524	163	687	0.24
	Unsafe	822	114	936	0.88
	Total	1,346	277	638	
Error of omission		0.61	0.59		0.39

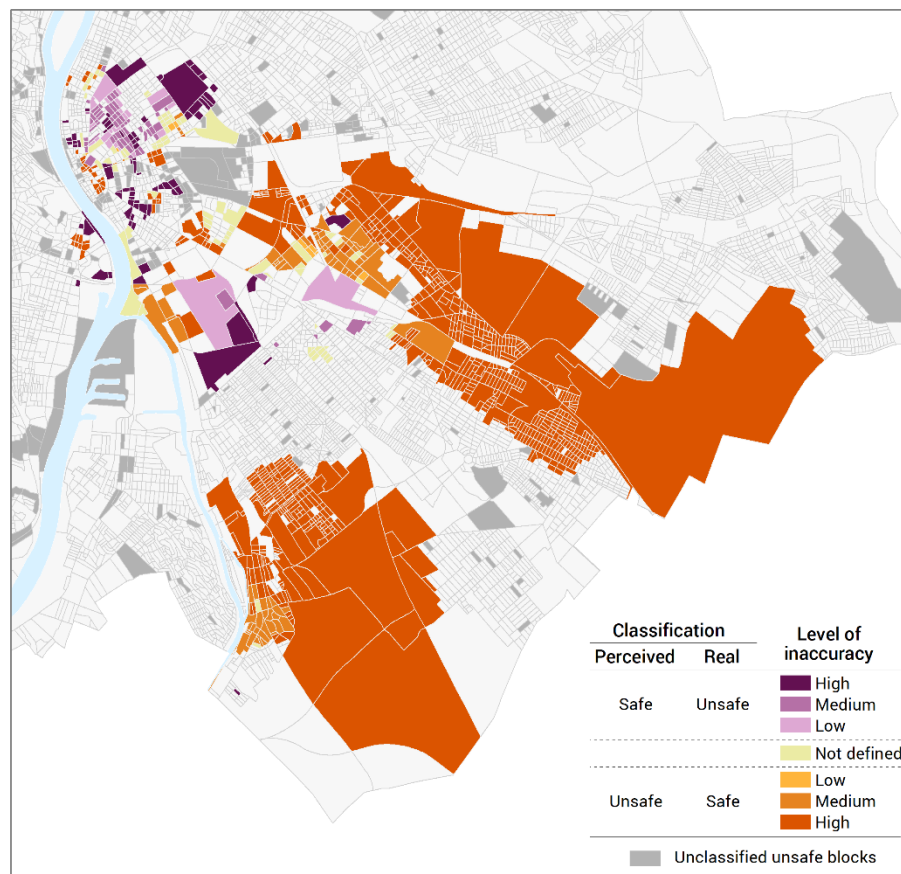


Fig. 6. Level of inaccurate perception of safe (IS) and unsafe (IU) areas per block.

5 Spatial Statistical Analysis of Perception Gaps

5.1 Methodology

The aim is to define if there is a relation between the locations of IS and the surrounding high crime rate areas or IU and the surrounding low crime rate. The analysis is not meant to explain the inaccuracy of perception but it will show the spatial relations between the two input variables. To address this aim, we use the bivariate local Moran's I, which is a spatial association measurement that relates the value of one variable in a given location and the average value of the neighboring features of a second variable. This means that the two variables are not analyzed in the same location. The value of the first variable in one location is compared with the average value of a conditional permutation performed with the neighboring features. The output is a cluster map which classifies the significant spatial units into high-high, low-low, high-low and low-high, where the first attribute corresponds to the value of the first variable and the second the value of the second variable in the neighboring areas.

The bivariate local Moran's I analysis was performed with a queen contiguity of first order and 999 permutations. The two input variables were the perceived classification given by the percentage of participants who identified a block as unsafe (>50% = unsafe area and <50% = safe area) and the number of events in the surrounding blocks.

5.2 Bivariate spatial autocorrelation analysis

The output map of the local Moran's I is shown in Figure 7. The turquoise color represents the perceived safe blocks, the dark ones are surrounded by blocks with low crime incidences, and the light ones are bounded by blocks with high crime incidences. Meanwhile, the brown blocks are perceived unsafe areas and contrary to the turquoise blocks, the neighboring blocks of the dark browns have high crime incidences and the light browns low crime incidences. Additionally, the hotspot blocks are shown for a better reference to the relationship between both variables. The light grey blocks are areas that are not significant, which means, those are blocks in which neighbors' values are not significantly different from the value resultant from a random permutation. The dark grey areas are those blocks that were not classified by the participants.

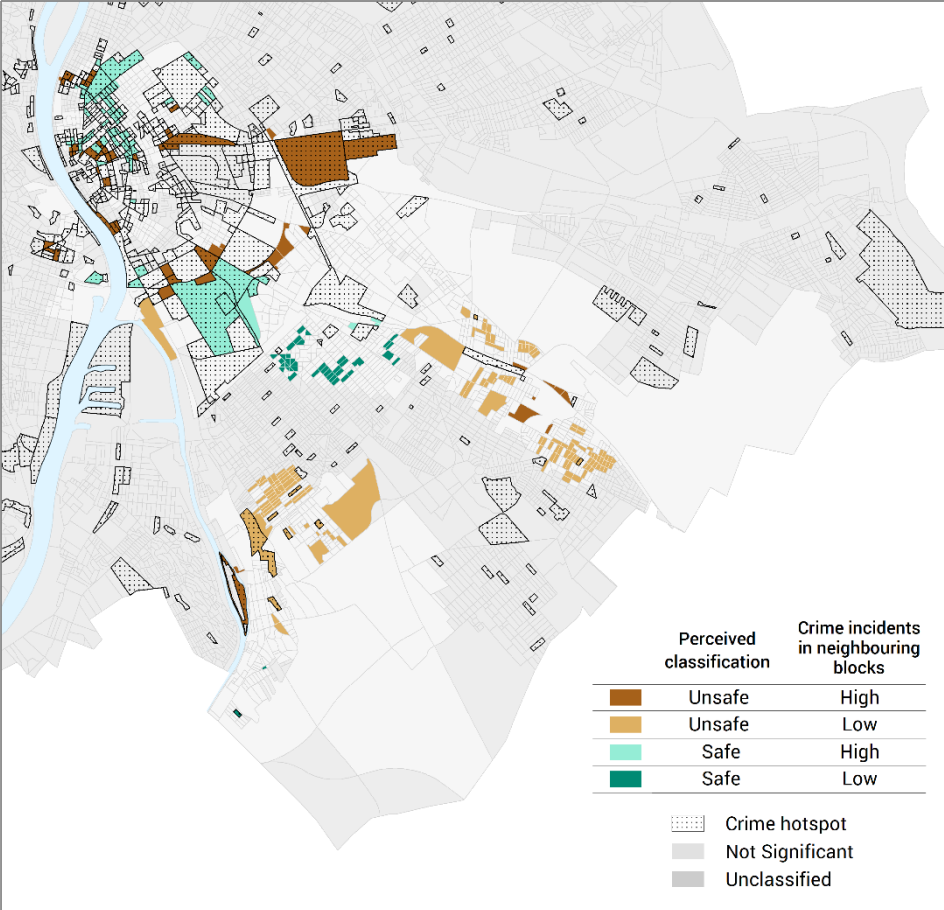


Fig. 7. Cluster map derived from a Bivariate Local Moran’s I analysis performed with the perceived classification in the target block and the number of crime incidents in the neighboring blocks.

The following step was to select, from the significant identified blocks in the bivariate spatial autocorrelation analysis (turquoise and brown blocks in Figure 7), those which were previously labeled as “inaccurately perceived” (Figure 6). Figure 8 shows the result of the selection. This map depicts in green those blocks that were inaccurately perceived as unsafe and of which neighboring blocks have high crime incidences. This relation could explain the inaccurate perception of safe areas, as the surroundings of the blocks perceived as unsafe could have an impact on people’s perception. They could believe, by spatial association that those selected blocks were actually unsafe areas due to the characteristics of the enclosing blocks. The red block was inaccurately perceived as a safe area, whereas it is unsafe in reality. Similarly to the previous case, this could be explained by the fact that the surrounding areas have low crime incidents and that due to the closest distance to low crime areas, the block is perceived as safe.

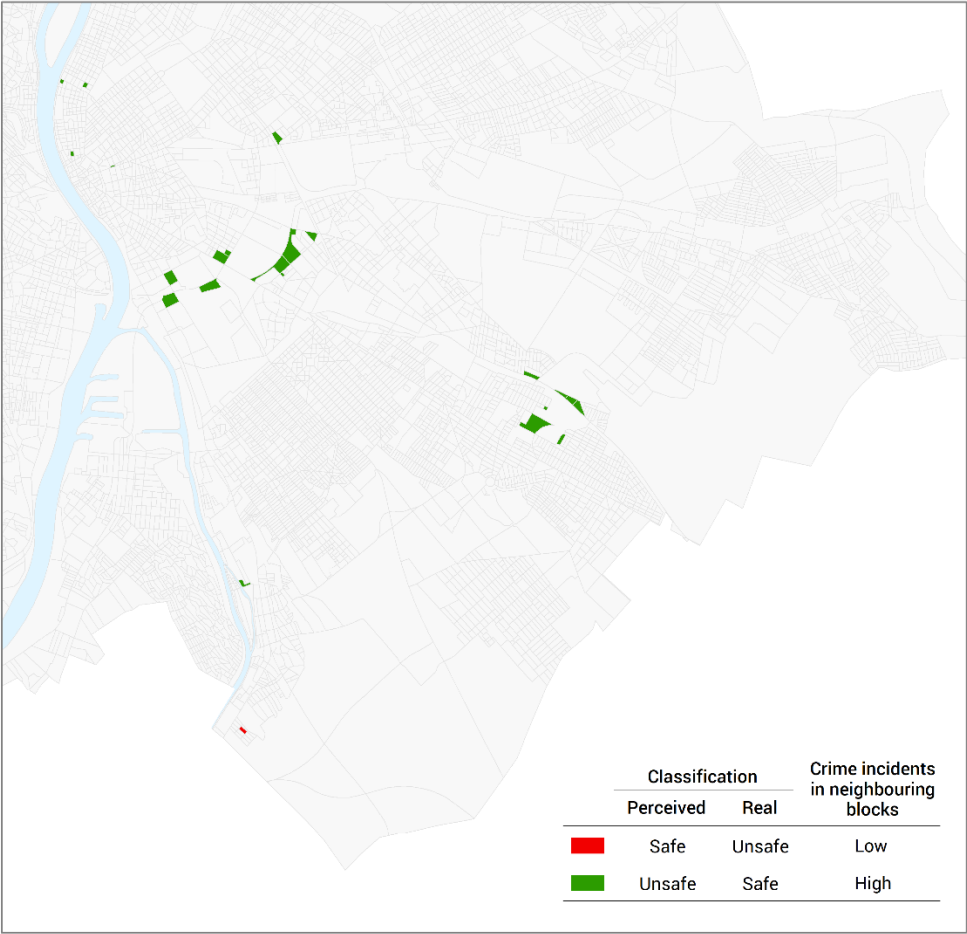


Fig. 8. Inaccurately perceived blocks colored by actual crime rate of their neighboring blocks.

6 Discussion

The first objective of this study was to analyze the location of perceived unsafe areas in relation to the distribution of crime incidents and people’s activity spaces. The participants identified safe areas closer to their neighborhood, which can be explained by the “endowment effect” [11]. Thus, people tend to have a perceptual bias due to a feeling of attachment towards their own community or neighborhood and value them “better” (i.e. more safe). Also, a higher percentage of participants identified unsafe areas further away from their daily routes. This result endorses the “geometry of crime” and “crime pattern” theories which as diverted by Spicer, Song and Brantingham [6], people would design daily routes through which they can stay off situations and places where they perceive as unsafe. Furthermore, this analytical part indicated perceptual gaps regarding the real spatial distribution of crime, which were analyzed in the second objective.

In addressing the second objective, accuracy was defined with four classes (i.e. AS, AU, IS, IU) and allows determining the blocks that could be priority areas for strategies

directed to narrow the perception gap. Besides the safe and unsafe places that were inaccurately classified (61% of the total classification), there are also those that were not considered by the participants and therefore also not in the analysis. Nevertheless, they must be taken into account for design of strategies as people are not aware of them. Furthermore, from a criminology theory perspective we should further explore if and how the crime perception gap in space affects the current crime prevalence of an area.

The third objective was to identify the relation between the location of the perception gap and the number of crimes reported in the surrounding areas. This type of analysis is suitable to explore the spatial association that people tend to do by transferring attributes from one location to adjacent areas. In our results, we identified that significant spatial patterns might have a *diffusion effect* on people's perception. For example, a safe place could be perceived as unsafe because the neighboring places are significant crime hotspots.

7 Future work

During the research process some difficulties arose that lead to *recommendations for future research*. First, the perception data collection by sketch maps must include a questionnaire or a think-aloud process to provide more information to the interpretation of the map. Although, the analysis of the data extracted can reveal valuable information, having an additional context of the participants' cognitive map can add more variables to explore that would lead to a better characterization of the people's perception.

Special attention should be paid on the processing and transformation processes, which can bring a certain level of uncertainty into the results. For example, in a point to polygon operation one should decide how to deal with a point located on the boundary shared by two polygons. Also, how to justify the size of these polygons (i.e. analysis units). Another case is the quality of the geocoding process. One issue found was the accuracy of the geocoding results, as for some records the points were located in the same pair of coordinates for similar but not same addresses.

Furthermore, data have been collected, processed, and analyzed using a polygon shape (i.e. areas). We used polygons as our representations because they were used as well in previous studies, as it is being discussed by Curtis [2]. However, areas are not homogeneous "realities" and an analysis based on the street network may reveal hidden variations (e.g., small streets being perceived differently than big ones).

Last, exploratory modelling can provide more input on the drivers of perceptual gaps by exploring additional spatial (or non-spatial) variables, for instance the land use or the average income. Although there are theories that explain the factors that sway the perception of crime, each city has different social dynamics where those factors may not have the same impact. Therefore, it is necessary to explore them to get a more precise overview of the context to be examined.

References

1. Sloan, N., Doran, B., Markham, F., & Pammer, K. (2016). Does base map size and imagery matter in sketch mapping?. *Applied geography*, 71, 24-31.
2. Curtis, J. W. (2012). Integrating sketch maps with GIS to explore fear of crime in the urban environment: A review of the past and prospects for the future. *Cartography and Geographic Information Science*, 39(4), 175-186.
3. Kohm, S. A. (2009). Spatial dimensions of fear in a high-crime community: Fear of crime or fear of disorder? *Canadian Journal of Criminology and Criminal Justice*, 51(1), 1-30.
4. Lopez, N., & Lukinbeal, C. (2010). Comparing Police and Residents' Perceptions of Crime in a Phoenix Neighborhood using Mental Maps in GIS. *Yearbook of the Association of Pacific Coast Geographers*, 33-55.
5. Lewis, D. A., & Maxfield, M. G. (1980). Fear in the neighborhoods: An investigation of the impact of crime. *Journal of research in crime and delinquency*, 17(2), 160-189.
6. Spicer, V., Song, J., & Brantingham, P. (2014). Bridging the perceptual gap: variations in crime perception of businesses at the neighborhood level. *Security Informatics*, 3(1), 14.
7. Foster, S., Giles-Corti, B., & Knuiman, M. (2010). Neighbourhood design and fear of crime: A social-ecological examination of the correlates of residents' fear in new suburban housing developments. *Health and Place*, 16(6), 1156-1165.
8. Fuhrmann, S., Huynh, N. T., & Scholz, R. (2013). Comparing fear of crime and crime statistics on a university campus. In *Crime Modeling and Mapping Using Geospatial Technologies* (pp. 319-337). Springer, Dordrecht.
9. Solymosi, R., Bowers, K., & Fujiyama, T. (2015). Mapping fear of crime as a context-dependent everyday experience that varies in space and time. *Legal and Criminological Psychology*, 20(2), 193-211.
10. Mohan, J., Twigg, L., & Taylor, J. (2011). Mind the double gap: Using multivariate multilevel modelling to investigate public perceptions of crime trends. *The British Journal of Criminology*, 51(6), 1035-1053.
11. Lora, E. (2016). The Distance between Perception and Reality in the Social Domains of Life. In *Handbook of Happiness Research in Latin America*, pp. 531-555. Springer, Dordrecht.
12. Duffy, B., Wake, R., Burrows, T. and & Bremer, P. (2008). Closing the gaps – crime and public perceptions. *International Review of Law Computers & Technology*, 22 (1-2), pp. 17-44.
13. Ardanaz, M., Corbacho, A., Ibarraran, P., & Ruiz-Vega, M. (2013). Mind the gap: Bridging perception and reality with crime information. *Unpublished Research Report. Washington, DC, United States: Inter-American Development Bank*.
14. Foster, S., Knuiman, M., Wood, L., & Giles-Corti, B. (2013). Suburban neighbourhood design: Associations with fear of crime versus perceived crime risk. *Journal of Environmental Psychology*, 36, 112-117.
15. Doran, B. J., & Burgess, M. B. (2011). Putting fear of crime on the map: Investigating perceptions of crime using geographic information systems. Springer Science & Business Media.
16. Cordner, G. (2010). *Reducing Fear of Crime Strategies for Police*. U.S. Department of Justice Office of Community Oriented Policing Services.
17. Tobler W., (1970) "A computer movie simulating urban growth in the Detroit region". *Economic Geography*, 46(Supplement): 234-240.
18. Póddör, A., Révész, A., Rácskai, P., & Sasvár, Z. Measuring citizens' fear of crime using a web application: A case study. In *GI Forum*, pp. 12-29, Salzburg, 2016.
19. Ratcliffe, J. H. (2004). Geocoding crime and a first estimate of a minimum acceptable hit rate. *International Journal of Geographical Information Science*, 18(1), 61-72.
20. Anselin, L. (1995). Local indicators of spatial association—LISA. *Geographical analysis*, 27(2), 93-115.

AGILE: GIScience Series, 1, 2020.

Full paper Proceedings of the 23rd AGILE Conference on Geographic Information Science, 2020.

Editors: Panagiotis Partinevelos, Phaedon Kyriakidis, and Marinos Kavouras

This contribution underwent peer review based on a full paper submission.

<https://doi.org/10.5194/agile-giss-1-20-2020> | © Authors 2020. CC BY 4.0 License.