Validation of a Google Street View-Based Neighborhood Disorder Observational Scale

Miriam Marco · Enrique Gracia ·
Manuel Martín-Fernández · Antonio López-Quílez

Published online: 24 February 2017
© The New York Academy of Medicine 2017

Abstract Recently, there has been a growing interest in developing new tools to measure neighborhood features using the benefits of emerging technologies. This study aimed to assess the psychometric properties of a neighborhood disorder observational scale using Google Street View (GSV). Two groups of raters conducted virtual audits of neighborhood disorder on all census block groups ($N=92$) in a district of the city of Valencia (Spain). Four different analyses were conducted to validate the instrument. First, inter-rater reliability was assessed through intraclass correlation coefficients, indicating moderated levels of agreement among raters. Second, confirmatory factor analyses were performed to test the latent structure of the scale. A bifactor solution was proposed, comprising a general factor (general neighborhood disorder) and two specific factors (physical disorder and physical decay). Third, the virtual audit scores were assessed with the physical audit scores, showing a positive relationship between both audit methods. In addition, correlations between the factor scores and socioeconomic and criminality indicators were assessed. Finally, we analyzed the spatial autocorrelation of the scale factors, and two fully Bayesian spatial regression models were run to study the influence of these factors on drug-related police interventions and interventions with young offenders. All these indicators showed an association with the general neighborhood

disorder. Taking together, results suggest that the GSV-based neighborhood disorder scale is a reliable, concise, and valid instrument to assess neighborhood disorder using new technologies.

Keywords Google Street View · Physical disorder · Physical decay · Neighborhood disorder · Virtual audits

Introduction

Neighborhood disorder has been related to a wide number of health outcomes such as obesity, chronic diseases, mortality, and sexually transmitted diseases [1–4], as well as to mental and behavioral disorders such as depression, anxiety, well-being, and binge drinking [5, 6]. Research has also found an association between neighborhood disorder and crimes such as juvenile delinquency, child maltreatment, and intimate partner violence [7–12].

To assess neighborhood disorder and related outcomes, researchers have traditionally used self-report measures [13–15] or physical systematic social observation [16–18]. Regarding the first, residents provide information about their own neighborhoods, which provide subjective data that it is difficult to obtain otherwise such as feelings of mistrust or fear, perceived safety, or social disorder [14, 15]. However, these data collection have some drawbacks, including that perceived neighborhood disorder may not match with the realities of the ecological context [19, 20], and neighbors may be influenced by stereotypes, personal characteristics, or the tendency to offer socially desirable responses [20].

M. Marco (✉) · E. Gracia · M. Martín-Fernández ·
A. López-Quílez
University of Valencia, Valencia, Spain
e-mail: Miriam.Marco-Francisco@uv.es

Regarding the physical systematic social observation, it overcomes some of these concerns, but this method has some major disadvantages: raters have to move to the study area (a very large area in some cases), and human resources and travel costs can be considerable [21–25]. In addition, conducting physical audits in highly conflictive areas that are difficult to access may pose dangers to research staff [24, 26].

These limitations have led to the development in recent years of innovative instruments to evaluate neighborhood characteristics using new technologies. One of the main alternatives is Google Street View, which has been used in a growing number of studies to assess neighborhood features [26–29]. Google Street View is freely available from the Google Maps website, and its database has images of almost all areas of western countries. Raters can visit neighborhoods and streets virtually, and due to its powerful zoom and high resolution, they can capture neighborhood characteristics in 360° images (Fig. 1). Google Street View is user-friendly, and it does not require any special computer skills.

The general aim of this study is to assess the psychometric properties of a neighborhood disorder observational scale using Google Street View in a Spanish city. The specific objectives of the study are as follows: (1) to assess the inter-rater reliability; (2) to test the latent structure of the scale through confirmatory factor analysis; (3) to assess the relationship between the virtual and the physical audit

scores; and (4) to assess the relation between the scale and some socioeconomic and criminality indicators.

Methods

Sample

The study area was the Maritime district of the city of Valencia, the third largest city in Spain. The city covers an area of 134.65 km² and has a population of 786,189 inhabitants (2015 data) [30]. The Maritime district is located in the east of the city, near the Mediterranean coast, and comprises very different neighborhoods in sociodemographic and physical terms. This district has a population of 57,710 inhabitants (2015 data) [30], and is divided into 92 census block groups, the smallest administrative unit available. We used census block groups as the neighborhood proxies. The virtual audits were conducted by two groups of raters using Google Street View imagery. Each group assessed the 92 census block groups.

Instrument

Google Street View (GSV)-Based Neighborhood Disorder Observational Scale We used a neighborhood disorder observational scale previously validated in the same city [31]. The original scale has 20 items and measures three factors: physical disorder, social disorder, and physical decay. The social disorder factor includes items such as people drinking alcohol or arguing in the street, and therefore could not be assessed with Google Street View imagery, which does not capture social interactions. This factor was therefore removed from the scale, leaving the two factors of physical disorder and physical decay. The physical disorder factor evaluates the presence of cigarette butts, trash, and empty bottles in the street as well as graffiti, including political and protest graffiti. Indicators of abandoned cars, used condoms, and syringes were removed from the scale because of their extremely low frequency. The physical decay factor assesses the presence of vacant or abandoned houses, abandoned commercial buildings, vandalized and run-down buildings, deteriorated residential units, and deteriorated recreation places. The final scale was composed of 10 Likert-type items ranging from 0 (no presence of the item) to 4 (high presence of the item) (see Appendix Table 5). The data were

FIG. 1 Google Street View image (above) vs. real photograph (below) of a boarded-up, vandalized house.

collected in 2015. All Google Street View images assessed were from 2014 and 2015.

Criterion-Related Validity Measures

Physical audits of neighborhood disorder in the same area were obtained to study criterion-related validity. We used physical audits conducted in a previous study in the same city area [31]. These audits were collected in 2014 by trained raters walking each census block group.

Two neighborhood characteristics frequently associated with neighborhood disorder were also assessed. Firstly, socioeconomic status was assessed by two indicators provided by the city's Statistics Office: education level (measured as the mean education level in the census block group on a four-point scale, where 1 = less than primary education, 2 = primary education, 3 = secondary education, and 4 = college education) and cadastral property value (average value of housing per square meter). Secondly, criminality indicators were used. To this end, police officers provided information about two types of police interventions: drugs-related interventions (police interventions related to possession, misuse, or distribution of drugs) and interventions with young offenders (police interventions where minors participate in illegal or disruptive behavior). Police officers' perceptions of the level of policing activity in each census block group was measured on two five-point scales evaluating the level of police intervention involving drugs or minors, respectively, where 0 = very low level of intervention and 4 = very high level of intervention.

Procedure

Two groups of two research assistants not directly involved in the project conducted virtual audits of neighborhood disorder in all the census block groups. In an initial training session led by the research staff, raters were introduced and instructed how to complete the GSV-based neighborhood disorder observational scale, and several examples were performed. Raters virtually walked around the whole census block group, and they assessed all the streets within. If the boundary divided a street or avenue into two, only the side corresponding to the census block group was assessed. Virtual audits were collected in 2015.

Data Analysis

Several analyses were performed to assess the psychometric properties of the scale. Inter-rater reliability was evaluated first using the intraclass correlation coefficient (ICC), computing the raw scores of each of the originally proposed factors of the GSV-based neighborhood disorder observational scale. Descriptive analyses were obtained for all items, as well as item-total corrected correlations.

To examine the latent structure of the scale, a confirmatory factor analysis was performed; various models were tested to achieve the best fit for the data. Three models were specified: a one-dimensional model, grouping all items into a general measure of neighborhood disorder; a two-dimensional model, based on the model proposed by Marco et al. [31], where social disorder items were removed, and only two correlated factors were specified (physical disorder and physical decay); and finally, a bifactor solution, which considered the previous two specific dimensions and also one general factor (i.e., general neighborhood disorder). Factors in this model are orthogonal, thus the general factor captured all the variance shared by the specific factors; each item loaded on one specific dimension and also on the general one. Given the categorical nature of the data, the WLSMV estimator was used to estimate the model parameters [32]. The fit of the models was assessed with the CFI and the TLI indices, following the criteria proposed by Hu and Bentler (CFI/TLI > 0.95) [33]. The RMSEA was also used to evaluate the residuals of the models. Values of the RMSEA less than 0.06 and 0.08 indicate good and mediocre fit, respectively [34]. Composite reliability index (CRI) also assessed reliability in conditions of non tau-equivalence [35].

To assess criterion-related validity, factor scores of the GSV-based neighborhood disorder observational scale were correlated with the criterion-related measures (physical audits, socioeconomic indicators, and criminality indicators). In addition, a spatial analysis was performed. Some studies have shown that neighborhood disorder tends to cluster spatially [36, 37]. We tested the presence of spatial autocorrelation using exploratory spatial data analysis with Moran's I index [38]. This statistic is commonly used to measure spatial autocorrelation, where positive values indicate the dataset exhibits a clustered pattern, negative values indicate the dataset displays a dispersed pattern, and zero indicates that there is no spatial clustering.

Finally, Bayesian spatial regression models were tested to analyze the influence of GSV-based neighborhood disorder observational scale factors in explaining drug-

Table 1 Item descriptive statistics and inter-rater reliability

Item	<i>M</i>	SD	Max	Min	r_{drop}		
Cigarettes	1.30	1.27	0	4	0.45		
Trash	1.55	1.27	0	4	0.63		
Bottles	0.55	0.94	0	4	0.72		
Graffiti	2.53	1.06	0	4	0.40		
Political graffiti	0.29	0.66	0	3	0.43		
Vacant houses	0.88	1.11	0	4	0.68		
Deteriorated commercial buildings	1.04	1.13	0	4	0.64		
Vandalized buildings	0.85	1.08	0	4	0.63		
Residential deterioration	1.16	1.28	0	4	0.70		
Deteriorated facilities	0.43	0.79	0	4	0.73		
Subscale	<i>M</i>	SD	Max	Min	ICC ₂	ICC _{2_k}	
Physical disorder	3.82	2.85	0	15	0.28	0.43	
Physical decay	6.24	3.67	0	16	0.38	0.55	

related police interventions and interventions with young offenders. We ran two spatial regression models using GSV-based neighborhood disorder observational scale scores as covariates to explain drug-related police interventions and interventions with young offenders, respectively. To measure the spatial component, we introduced two random effects: a structured spatial effect S , and an unstructured spatial effect U to assess heterogeneity. All analyses were performed with the statistical software *R* except the confirmatory factor analyses, which were performed with *MPlus 7* [39, 40].

Results

Descriptive Statistics and Inter-Rater Reliability

The mean, standard deviation, and range of each item are displayed in Table 1. Descriptive analysis showed that the items are slightly displaced to the left (mean range 0.43–2.53 with a mean standard deviation around 1), indicating, on average, a low presence of the observed indicators of physical disorder and decay. The item-total corrected correlations were moderately high, ranging from 0.40 to 0.73, which suggests that the items were related to the measured construct. Intraclass correlation coefficients for random raters (ICC₂) and average intraclass correlation coefficients (ICC_{2_k}) were between the standard cut-offs, indicating a fair agreement if all the items are taken into account (between 0.21 and 0.40), and moderate agreement if only the mean of each subscale is considered (between 0.41 and 0.60) [41].

Factor Structure and Scale Reliability

Three models were fitted and compared to evaluate the latent structure of the scale.

As shown in Table 2, the three models yielded adequate relative goodness-of-fit indices. Though the relative fit indices were fair for the one-dimensional and the two-dimensional models, the residuals for both models indicated a poor fit to the data. The bifactor solution had an excellent fit in both indices, however, and the third model was therefore retained.

Figure 2 includes the model diagram with the standardized parameter estimates. All items loaded significantly on the general factor, whereas loadings for some of them were not significant on their specific factors. Specifically, indicators of political or protest graffiti, abandoned commercial buildings, and deteriorated recreation places loaded only on the general dimension. Taking the bifactor model as the latent structure of the scale, the CRI was obtained. The CRI for the general neighborhood disorder dimension was very high (CRI = 0.93), although the reliability was lower for the specific dimensions of physical disorder (CRI = 0.57) and physical decay (CRI = 0.61).

Criterion-Related Validity

GSV-based neighborhood disorder observational scale scores were related to the criterion measures (see Table 3). First, the factor scores of the physical and virtual audits were correlated. Note that the physical audit scores followed the original model proposed by

Table 2 CFA fit indices

Model	χ^2	df	CFI	TLI	RMSEA	RMSEA ₀₅	RMSEA ₉₅
One Factor	86.412	35	0.950	0.936	0.126	0.093	0.160
Two Factor	74.656	34	0.961	0.948	0.114	0.079	0.149
Bifactor	33.430	59	0.992	0.985	0.061	0.000	0.110

CFA confirmatory factor analysis

Marco et al. [31], whereas the virtual audits followed the bifactor model. Scores obtained with the GSV-based neighborhood disorder observational scale were positively related to the physical audit scores. The general neighborhood disorder factor of the virtual audits showed an extremely high relation between the physical audit factors (physical disorder, $r_{xy} = 0.95$, $t(90) = 27.97$, $p < 0.001$, and physical decay, $r_{xy} = 0.97$, $t(90) = 38.27$, $p < 0.001$), meaning that the virtual audit scores ordered the census blocks groups in a similar way to the physical audit scores. Regarding the specific factors of the virtual audits, both physical disorder and physical decay scores were positively related with the physical audits of disorder, $r_{xy} = 0.39$, $t(90) = 3.98$, $p < 0.001$, and decay, $r_{xy} = 0.37$, $t(90) = 3.74$, $p < 0.001$.

On the other hand, the general neighborhood disorder factor was highly correlated with socioeconomic indicators: higher scores in the general factor were related to lower neighborhood education level, $r_{xy} = -0.33$, $t(90) = -3.48$, $p < 0.001$, and lower cadastral property value, $r_{xy} = -0.34$, $t(90) = -3.37$, $p < 0.001$. In addition,

the general neighborhood disorder factor was strongly correlated with police interventions, especially those involving young offenders, $r_{xy} = 0.40$, $t(90) = 4.15$, $p < 0.001$, but also drug-related interventions, $r_{xy} = 0.33$, $t(90) = 3.27$, $p < 0.001$. The specific factors, however, showed no significant association with criminality indicators, and only the physical decay factor was negatively related to cadastral property value, $r_{xy} = -0.28$, $t(90) = -2.80$, $p = 0.006$, and education level, $r_{xy} = -0.28$, $t(90) = -2.81$, $p = 0.006$. Therefore, high scores in the physical decay factor tended to co-occur with lower cadastral property values and lower education level.

In the spatial analysis, the three factors showed significant positive spatial autocorrelation (general neighborhood disorder: $I = 0.41$, $p < 0.001$; physical disorder: $I = 0.20$, $p < 0.001$; physical decay: $I = 0.30$, $p < 0.001$), indicating a tendency to cluster together geographically, rather than being randomly distributed in space. In addition, in the two spatial Bayesian regression models, general neighborhood disorder and physical decay

FIG. 2 CFA bifactor model. Factors and factor loadings of each item.

Table 3 Criterion-related validity correlations

	General neighborhood disorder	Physical disorder	Physical decay
Physical audits			
Physical disorder	0.95	0.39	0.12
Physical decay	0.97	0.14	0.37
Socioeconomic indicators			
Cadastral property value	−0.33	−0.11	−0.28
Education level	−0.34	−0.14	−0.28
Criminality indicators			
Drug-related police interventions	0.33	0.09	0.17
Young offenders police interventions	0.40	−0.09	0.19

showed more than 95% posterior probability of being higher than zero. These results indicated a higher prevalence of drug-related police interventions and interventions with young offenders in areas with higher general neighborhood disorder and physical decay. Physical disorder showed no relevant relationship with drug-related interventions nor with interventions with young offenders. Table 4 summarizes the models.

Discussion

The aim of this study was to develop and analyze the psychometric properties of a neighborhood observational scale using Google Street View. In light of the results obtained, we can conclude that the GSV-based neighborhood disorder observational scale provides an accurate, concise, and valid measure of neighborhood

disorder, comparable to measures obtained with traditional observational instruments. The strong relation between physical and virtual audits suggests it is feasible to conduct neighborhood disorder assessments using Google Street View.

We found that the bifactor model was the latent structure that best fitted the data of the models compared for the virtual audit of neighborhood disorder. This model yielded two specific dimensions that covered the unique aspects of physical disorder and physical decay. The model revealed also a general neighborhood disorder dimension that explained all the shared variance not captured by the specific dimensions [42]. Thus, the common elements of the specific factors originally proposed by Marco et al. [31] could be unified into one general neighborhood disorder dimension in the virtual neighborhood disorder audits.

The internal consistency of the general neighborhood disorder factor was very high, indicating a good overall

Table 4 Bayesian regression models with two dependent variables: drug-related police interventions and police interventions with young offenders

Explanatory variables	Drug-related police interventions			Police interventions with young offenders		
	Mean	Std. error	95% CrI ^a	Mean	Std. error	95% CrI
Intercept	0.401	0.061	0.275, 0.513	0.398	0.061	0.272, 0.511
General neighborhood disorder	0.210	0.091	0.034, 0.395	0.209	0.088	0.025, 0.375
Physical disorder	0.061	0.083	−0.097, 0.222	0.061	0.085	−0.094, 0.232
Physical decay	0.173	0.100	−0.030, 0.370	0.172	0.098	−0.015, 0.364
σ_s^b	0.201	0.126	0.017, 0.427	0.197	0.126	0.014, 0.446
σ_u^c	0.312	0.175	0.028, 0.660	0.310	0.178	0.013, 0.681

^a Credible interval

^b Standard deviation spatially structured term;

^c Standard deviation unstructured term

reliability of the GSV-based neighborhood disorder observational scale.

Several variables were used to study criterion-related validity. On the one hand, the strong relationship between the virtually audited general neighborhood disorder factor and the two physically audited factors indicated that the factor scores of the two methods ordered the census block groups almost identically. This result is not entirely unexpected if we take into account that the bifactor solution modeled the common elements (i.e., the relation) of the physical disorder and decay factors into one general factor. Therefore, a strong correlation would be expected between the physically audited factors—which followed a two correlated factor model—and a non-specific measure of the shared variance of the same factors in the virtual audit. Likewise, the correlations between the virtually audited specific factors (i.e., physical disorder and decay) and their physically audited counterparts reflected that the core and non-shared elements of both specific factors—those accounted for by the bifactor solution—were positively related with the factors found in the original scale [31]. These results suggest that, in general, virtual and physical audits of the characteristics of disorder tend to be very similar.

On the other hand, we found an association between the GSV-based neighborhood disorder observational scale and drug-related interventions and those involving young offenders. These results support previous research findings that neighborhoods with higher levels of disorder are likely to show higher crime rates [11, 12, 43]. Similarly, the relationship between neighborhood disorder and socioeconomic indicators are in line with previous studies linking neighborhood disorder and socioeconomic neighborhood characteristics [16, 22, 44, 45]. The association found between neighborhood disorder and criminality indicators suggests that this instrument may be useful for studying neighborhood criminality.

The spatial analyses showed that the three disorder factors (general neighborhood disorder, physical disorder, and physical decay) were spatially clustered as expected [36, 37]. We then conducted two spatial Bayesian regressions to study the two criminality indicators. We found that general neighborhood disorder and physical decay were related to both types of police interventions, but there was no such association with physical disorder. These results might be due to the fact that the general neighborhood disorder captures part of the variability of physical disorder, and this part would

explain the relationship between policing activity and disorder.

In recent years there has been growing interest in the use of computer applications to assess neighborhood characteristics. Some research has shown the advantages of using this kind of instrument [21, 22, 25, 27]. Specifically, previous research has shown Google Street View can be a useful research instrument [22, 28, 29]. Our results provide new evidence of the benefits of using Google Street View in applications such as neighborhood disorder assessment [21, 22, 29]. This measurement approach avoids researchers having to move around the city and face the risks of visiting potentially hazardous areas [26], and reduces observation time [27]. Nevertheless, some studies found that virtual audits were faster than physical audits [27, 46], while others showed opposite results [47]. Future studies would benefit from including observation times measures to assess the differences between physical and virtual audits. Also, most of the studies assessing neighborhood characteristics using Google Street View have been conducted in Anglo-Saxon cities [48, 49], and only few studies have been conducted in the context of European countries [47]. This study provides new evidence about virtual audits in a European city.

Among the limitations, the GSV-based neighborhood disorder scale can only obtain static observations, and generally Google images are of neighborhoods in the same or the previous year. Although Google Street View is constantly being updated, unlike traditional physical audits it can be more difficult to use for longitudinal studies. Some urban areas are assessed from Google Street View less frequently than others; also, Google Street View does not capture time periods shorter than a year. However, the virtual data have the potential to facilitate longitudinal studies if archived imagery is available. Another limitation is that some characteristics often studied in the neighborhood context cannot be assessed with this instrument because it is not accurate enough to observe small items such as syringes and condoms in the street [22, 24, 29], which we had to remove from the final analysis. Likewise, social disorder, which has typically been considered as another relevant dimension of neighborhood disorder [18, 50], cannot be observed through Google Street View because it refers to potentially threatening social interactions such as street fights, the presence of homeless and drunk people, street prostitution, etc. [18, 50]. However, Google Street View could be used in conjunction with other

measures such as resident surveys and physical audits to address these limitations.

In conclusion, this study provides evidence of a reliable and valid instrument to evaluate neighborhood disorder using new technologies. Google Street View can be a useful tool to avoid the problems of physical audits, and it can contribute to the study of various social and health outcomes.

Acknowledgments This research was supported by the Spanish Ministerio de Economía y Competitividad (grant number PSI2014-54561-P). Miriam Marco was supported by the FPU program of the Spanish Ministerio de Educación, Cultura y Deporte (FPU2013/00164). Manuel Martín- Fernández was supported by the FPI program of the Spanish Ministerio de Economía y Competitividad (BES-2015-075576). We wish to thank José Serrano, Chief of the Valencia city Police Department, for his support and assistance in collecting criminality data for this study.

Appendix

Table 5 GSV-based neighborhood disorder observational scale items

1. Cigarettes in the street
2. Trash in the street
3. Empty bottles in the street
4. Graffiti
5. Political or protest message graffiti
6. Vacant or abandoned houses
7. Abandoned commercial buildings
8. Vandalized and run-down buildings
9. Deteriorated residential units
10. Deteriorated recreation places

References

1. Feng J, Glass TA, Curriero FC, Stewart WF, Schwartz BS. The built environment and obesity: a systematic review of the epidemiologic evidence. *Health Place*. 2010; 16: 175–90.
2. Latkin CA, Curry AD, Hua W, Davey MA. Direct and indirect associations of neighborhood disorder with drug use and high-risk sexual partners. *Am J Prev Med*. 2007; 32(Suppl 6): S234–41.
3. Winkleby MA, Cubbin C. Influence of individual and neighbourhood socioeconomic status on mortality among Black, Mexican-American, and White women and men in the United States. *J Epidemiol Community Health*. 2003; 57: 444–52.
4. Rose G. Sick individuals and sick populations. *Int J Epidemiol*. 1984; 14: 32–8.
5. Hill TD, Angel RJ. Neighborhood disorder, psychological distress, and heavy drinking. *Soc Sci Med*. 2005; 61: 965–75.
6. Keyes KM, McLaughlin KA, Koenen KC, Goldmann E, Uddin M, Galea S. Child maltreatment increases sensitivity to adverse social contexts: neighborhood physical disorder and incident binge drinking in Detroit. *Drug Alcohol Depend*. 2012; 122: 77–85.
7. Law J, Quick M. Exploring links between juvenile offenders and social disorganization at a large map scale: a Bayesian spatial modeling approach. *J Geogr Syst*. 2013; 15: 89–113.
8. Freisthler B, Bruce E, Needell B. Understanding the geospatial relationship of neighborhood characteristics and rates of maltreatment for Black, Hispanic, and White children. *Soc Work*. 2007; 52: 7–16.
9. Cunradi CB. Drinking level, neighborhood social disorder, and mutual intimate partner violence. *Alcohol Clin Exp Res*. 2007; 31: 1012–9.
10. Cunradi CB. Intimate partner violence among Hispanic men and women: the role of drinking, neighborhood disorder, and acculturation-related factors. *Violence Vict*. 2009; 24: 83–97.
11. Gracia E, López-Quílez A, Marco M, Lladosa S, Lila M. The spatial epidemiology of intimate partner violence: do neighborhoods matter? *Am J Epidemiol*. 2015; 182: 58–66.
12. Gracia E, López-Quílez A, Marco M, Lladosa S, Lila M. Exploring neighborhood influences on small-area variation in intimate partner violence risk: a Bayesian random-effect modeling approach. *Int J Environ Res Public Health*. 2014; 11: 866–82.
13. Gracia E, Fuentes MC, García F, Lila M. Perceived neighborhood violence, parenting styles, and developmental outcomes among Spanish adolescents. *J Community Psychol*. 2012; 40: 1004–21.
14. Ross CE, Mirowsky J. Neighborhood disadvantage, disorder, and health. *J Health Soc Behav*. 2001; 42: 258–76.
15. Ross CE, Mirowsky J. Neighborhood disorder, subjective alienation, and distress. *J Health Soc Behav*. 2009; 50: 49–64.
16. Sampson RJ, Raudenbush SW. Systematic social observation of public spaces: a new look at disorder in urban neighborhoods. *Am J Sociol*. 1999; 105: 603–51.
17. McDonell JR, Waters TJ. Construction and validation of an observational scale of neighborhood characteristics. *Soc Indic Res*. 2011; 104: 439–57.
18. Sampson RJ, Raudenbush SW. Seeing disorder: neighborhood stigma and the social construction of “broken windows”. *Soc Psychol Q*. 2004; 67: 319–42.
19. Gracia E, Herrero J. Perceived neighborhood social disorder and residents’ attitudes toward reporting child physical abuse. *Child Abuse Negl*. 2006; 30: 357–65.
20. Gracia E, Herrero J. Perceived neighborhood social disorder and attitudes toward reporting domestic violence against women. *J Interpers Violence*. 2007; 22: 737–52.
21. Bader MDM, Mooney SJ, Lee YJ, et al. Development and deployment of the computer assisted neighborhood visual assessment system (CANVAS) to measure health-related neighborhood conditions. *Health Place*. 2015; 31: 163–72.
22. Mooney SJ, Bader MDM, Lovasi GS, Neckerman KM, Teitler JO, Rundle AG. Validity of an ecometric neighborhood physical disorder measure constructed by virtual street audit. *Am J Epidemiol*. 2014; 180: 626–35.

23. Mooney SJ, DiMaggio CJ, Lovasi GS, et al. Use of Google Street View to assess environmental contributions to pedestrian injury. *Am J Public Health*. 2016; 106: 462–9.
24. Rundle AG, Bader MDM, Richards CA, Neckerman KM, Teitler JO. Using Google Street View to audit neighborhood environments. *Am J Prev Med*. 2011; 40: 94–100.
25. Wilson JS, Kelly CM, Schootman M, et al. Assessing the built environment using omnidirectional imagery. *Am J Prev Med*. 2012; 42: 193–9.
26. Griew P, Hillsdon M, Foster C, Coombes E, Jones A, Wilkinson P. Developing and testing a street audit tool using Google Street View to measure environmental supportiveness for physical activity. *Int J Behav Nutr Phys Act*. 2013; 10: 103–9.
27. Badland HM, Opit S, Witten K, Kearns RA, Mavoa S. Can virtual streetscape audits reliably replace physical streetscape audits? *J Urban Health*. 2010; 87: 1007–16.
28. Kelly CM, Wilson JS, Baker EA, Miller DK, Schootman M. Using Google Street View to audit the built environment: inter-rater reliability results. *Ann Behav Med*. 2013; 45(Suppl 1): S108–12.
29. Odgers CL, Caspi A, Bates CJ, Sampson RJ, Moffitt TE. Systematic social observation of children's neighborhoods using Google Street View: a reliable and cost-effective method. *J Child Psychol Psychiatry*. 2012; 53: 1009–17.
30. Statistics Office. *Padrón Municipal de Habitantes a 1 de Enero de 2015*. Valencia, Spain: Statistics Office, Valencia Town Hall; 2016.
31. Marco M, Gracia E, Tomás JM, López-Quílez A. Assessing neighborhood disorder: validation of a three-factor observational scale. *Eur J Psychol Appl Legal Context*. 2015; 7: 81–9.
32. Finney SJ, DiStefano C. Non-normal and categorical data in SEM. In: Hancock GR, Mueller RO, eds. *Structural Equation Modelling: a Second Course*. Greenwich, CO: Information Age Publishing; 2006: 269–314.
33. Hu L, Bentler PM. Cut-off criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. *Struct Equ Model*. 1999; 6: 1–55.
34. MacCallum RC, Browne MW, Sugawara HM. Power analysis and determination of sample size for covariance structure modeling. *Psychol Methods*. 1996; 1: 130–49.
35. Graham JM. Congeneric and essentially tau-equivalent estimates of score reliability: what they are and how to use them. *Educ Psychol Meas*. 2006; 66: 930–44.
36. Bruisma GJN, Pauwels LJR, Weerman FM, Bernasco W. Social disorganization, social capital, collective efficacy and the spatial distribution of crime and offenders. *Br J Criminol*. 2013; 53: 942–63.
37. DiMaggio C. Small-area spatiotemporal analysis of pedestrian and bicyclist injuries in New York City. *Epidemiology*. 2015; 26: 247–54.
38. Anselin L. *Spatial Econometrics, Methods and Models*. Dordrecht, the Netherlands: Kluwer Academic; 1988.
39. R Core Team. *R: a Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing; 2015.
40. Muthén LK, Muthén BO. *Mplus User's Guide (sixth edition)*. Los Angeles, CA: Muthén & Muthén; 2007.
41. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics*. 1977; 33: 159–74.
42. Reise SP. The rediscovery of bifactor measurement models. *Multivar Behav Res*. 2012; 47: 667–96.
43. Lum C. Violence, drug markets and racial composition: challenging stereotypes through spatial analysis. *Urban Stud*. 2011; 48: 2715–32.
44. Caughy MO, O'Campo PJ, Patterson J. A brief observational measure for urban neighborhoods. *Health Place*. 2001; 7: 225–36.
45. Sampson RJ, Raudenbush SW, Earls F. Neighborhoods and violent crime: a multilevel study of collective efficacy. *Science*. 1997; 277: 918–24.
46. Bethlehem JR, Mackenbach JD, Ben-Rebah M, et al. The SPOTLIGHT virtual audit tool: a valid and reliable tool to assess obesogenic characteristics of the built environment. *Int J Health Geogr*. 2014; 13: 52.
47. Gullón P, Badland HM, Alfayate S, et al. Assessing walking and cycling environments in the streets of Madrid: comparing on-field and virtual audits. *J Urban Health*. 2015; 92: 923–39.
48. Le Galès P, Zagrodzki M. *Cities are Back in Town: the US/Europe Comparison*. Cahier Européen numéro 05/06 du Pôle Ville/métropolis/cosmopolis. Paris, France: Centre d'Etudes Européennes de Sciences Po; 2006.
49. Summers AA, Cheshire PC, Senn L. *Urban Change in the United States and Western Europe: Comparative Analysis and Policy*. Washington, DC: The Urban Institute Press; 1999.
50. Sampson RJ. Disparity and diversity in the contemporary city: social (dis)order revisited. *Br J Sociol*. 2009; 60: 1–31.