Cite as (APA 7th):

Hardyns, W., Snaphaan, T., Willems, S., & Pauwels, L. J. R. (2021). Collective efficacy and disorder through the eyes of neighbourhood inhabitants and key informants. *European Journal of Criminology*. Advance online publication.

Collective efficacy and disorder through the eyes of neighbourhood inhabitants and key informants

Wim Hardyns[1], Thom Snaphaan[1], Sara Willems[2] & Lieven J. R. Pauwels[1]

[1] Department of Criminology, Criminal Law and Social Law, Faculty of Law and Criminology, Ghent University, 9000 Ghent, Belgium.

[2] Department of Public Health and Primary Care, Faculty of Medicine and Health Sciences, Ghent University, 9000 Ghent, Belgium.

Abstract

This study examines the ecological reliability, convergent validity and ecological stability of neighbourhood (dis)organisational processes measured by means of two methods: inhabitant surveys and the so-called key informant analysis technique. Considering that ecological processes play a major role in many contemporary criminological theories and research, it is vital to take into account methodological challenges and question the reliability, validity and stability of the measures reflecting these underlying processes. (Dis)organisational processes are predominantly measured by means of questionnaires probed to neighbourhood inhabitants. This approach requires large numbers of respondents to yield ecologically reliable and valid measures. In this study we analyse the relationships between ecological measures of neighbourhood processes based on surveys of inhabitants versus key informants. The findings suggest that key informants can provide reliable, valid and stable measures of (dis)organisational neighbourhood processes. Therefore, the key informant analysis technique is an essential complementary, or even substitutive, method in the measurement of neighbourhood processes; shared survey-method variance is eliminated and it is possible to survey fewer key informants than inhabitants to obtain reliable and valid information on social trust and disorder. Nevertheless, this method is not suitable for measuring all neighbourhood processes, such as informal social control. Therefore, outstanding challenges and avenues for future research are discussed as well.

Keywords

Collective efficacy, disorder, inhabitant survey, key informant analysis, neighbourhood, social disorganisation

Introduction

Ecological processes have been studied in urban sociology and criminology from the early days of the Chicago School. Scholars like Park and Burgess (1929), and Shaw and McKay (1942) pointed to the importance of studying differences and changes in community social structural characteristics. Especially structural disadvantage (e.g., concentrated poverty), social organisational processes (e.g., social trust, informal social control) and disorganisational processes (e.g., disorder) as neighbourhood characteristics received attention from scholars since then (Bruinsma and Johnson, 2018; Bursik and Grasmick, 1993; Kubrin and Wo, 2016; Kurtz, 1984; Park et al., 1967 [1925]; Sampson et al., 1997; Steenbeek, 2011). From a theoretical point of view these ecological processes refer to the social organisation of a community, rather than the organisation of individuals embedded in those communities. Social trust, informal social control, and both physical and social disorder have evolved from theoretical constructs to operationalized constructs in ecological studies since the 1980s (Sampson and Groves, 1989). Nevertheless, Groff (2018: 115) recently states that "the inclusion of collective social process is underdeveloped and reflects the difficulty in developing measures of those processes".

Today, ecological processes have been integrated in cross-level theories and inquiries on the role of ecological processes on delinquency, victimization, fear and many other outcomes (Wikström 2007). Such cross-level integrated theories address the finding that the relationship between structural factors and crime is not as straightforward as some scholars assume (e.g., Wikström and Treiber, 2016). Traditional social disorganisation theories are criticized for the conceptual limitation regarding the relative lack of attention paid to the mechanisms that mediate the effect of structural neighbourhood characteristics and, hence, "perhaps the greatest [challenge] involves identifying and measuring the social mechanisms that account for heightened crime rates in socially disorganized neighborhoods" (Kubrin and Wo, 2016: 123). Contemporary elaborations of these social disorganisation theories, like collective efficacy theory (Sampson 2012; Sampson et al., 1997) have addressed these shortcomings in several ways. Collective efficacy is considered as the key ecological mechanism that explains why some neighbourhoods with predisposing structural characteristics experience high levels of crime and disorder, whereas others do not (Brunton-Smith et al., 2018; Hardyns, Snaphaan, Pauwels, et al., 2019). Collective efficacy is considered to be an attribute of neighbourhoods rather than of individuals: a combination of the networks, norms, and trust between residents and the capacity this endows them with to control and suppress anti-social and criminal behaviour (Mazerolle et al., 2010; Sampson, 2012; Zhang et al., 2007). However, the collective and at the same time inherently subjective nature of collective efficacy as a concept poses challenges for valid and robust measurement (Brunton-Smith et al., 2018; Hipp, 2016). Given the major role that ecological processes play in many contemporary criminological theories and research, it is vital to take into account methodological challenges and question the reliability, validity and stability of the measures reflecting these underlying processes.

In this study we examine the ecological reliability, convergent validity and ecological stability of neighbourhood (dis)organisational processes, by comparing surveys of inhabitants to surveys of key informants. We focus on the measurement of social trust, informal social control and disorder as neighbourhood (dis)organisational processes.[1] The unique contribution of this study compared to prior research is provided by (1) the multi-method design and (2) the rather small units of analysis. First, prior research used either a quantitative approach with surveying neighbourhood inhabitants (e.g., Oberwittler and Wikström, 2009) or a qualitative approach with focus groups and/or the key informant analysis technique (e.g., Gerell, 2015). Second, in prior comparative research rather large units of analysis such as neighbourhood clusters or postcode areas have been used (e.g., Pauwels and Hardyns, 2009; around 10,000 to 16,000 inhabitants per ecological unit on average).

Measuring neighbourhood (dis)organisation processes

The most commonly used method to capture ecological processes is the survey of inhabitants of ecological entities, like neighbourhoods. However, one can argue that inhabitants do not necessarily need to be aware of social processes arising in their residential areas, as many of them commute and therefore do not necessarily have a clear idea of what is really going on in their residential areas. Using only residents as subjects may thus lead to the introduction of measurement error and bias. It is important, therefore, to develop alternative ways of measuring social processes. Oberwittler and Wikström (2009) demonstrated that smaller units of analysis generate more reliable ecological measures by surveying neighbourhood inhabitants. Earlier, Raudenbush and Sampson (1999) suggested that 20 to 30 respondents (inhabitants) can be sufficient to reliably measure neighbourhood social processes. They also found that more than

[1] Philosophical discussions on the extent to which ecological processes refer separate ontological neighbourhood characteristics should be studied as a topic of its own. Bunge's emergent systemism can be a theoretical guideline (see Bunge, 2003).

40 respondents provides little incremental improvement of ecological reliability. Raudenbush and Sampson used combined census tracts as their operational measure of local communities. It is therefore difficult to generalize their findings to other units of analysis that also refer to small areas. In general, the higher the level of analysis, the more heterogeneous the area (Oberwittler and Wikström, 2009). It is also reasonable to assume that with larger areas, more inhabitants are necessary to obtain reliable measures of processes at the ecological level. On the other hand, when ecological units get smaller, more units of the same larger entity (e.g. a city) are included. This altogether requires more respondents, when it is necessary to obtain a representative sample per ecological unit, and as a consequence survey costs may rise substantially.

One interesting and useful alternative to a survey of inhabitants is the use of systematic social observation (SSO), a technique that has successfully been used to measure both social and physical disorder (Mastrofski et al., 2010; Raudenbush and Sampson, 1999). SSO avoids bias in respondents' lack of knowledge on disorder and directly measures visible aspects of neighbourhood disorganisational processes, such as public alcohol consumption on streets, the presence of litter, and graffiti. The large advantage of using SSO as an additional data source is to overcome the problem of single-source bias (Campbell and Fiske, 1959; Thorndike, 1920) or shared survey-method variance (Sampson and Raudenbush, 1999; Taylor, 1999). However, this research method is also prone to several errors and biases (e.g., Hoeben et al., 2018). The high cost of this method is another reason why it has never (as far as we know) been used within European empirical research on area variation in crime and disorder on a scale comparable to the Chicago (PHDCN) study in the mid-90s (Earls et al., 2005). However, novel methods deliver opportunities in performing SSO in a reliable, valid and more cost efficient way (Kronkvist, 2013; Odgers et al., 2012). The method of SSO may be accurate when studying (visible signs of) disorder, but it may not capture social cohesive processes accurately.

One less recognized method to measure social processes in local areas such as neighbourhoods is key informant analysis (KIA). This method has been used in Swedish neighbourhood research (Tiby and Olsson, 1997). Pauwels and Hardyns (2009) demonstrated that the technique of KIA could be used to create ecologically reliable and valid measures of neighbourhood (dis)organisational processes, referring to social trust and disorder.

"Key informants are defined as persons that have a 'privileged' position to provide detailed information on local area processes" (Pauwels and Hardyns, 2009: 404). By privilege position, we refer to a position allowing the person to adequately judge the area social climate. Therefore key informants are expected to have above-average knowledge of ecological processes such as social trust, informal social control and disorder. As a consequence, fewer informants are necessary to get reliable measures of ecological processes. Some people can, through their social position or job, provide more meaningful and less biased information on these matters. For that reason, the selection process of key informants is of paramount importance.

Key informants that meet this criterion of above-average knowledge of local area processes were previously identified in jobs such as social work, local police, local shops (e.g., groceries, newspaper shops), local pubs and local policy work. These key informants can be given self-administered questionnaires, rather similar to conducting a survey of neighbourhood inhabitants. One major difference between the use of surveys of inhabitants and profession-based key informants is the selection procedure employed. While random selection is the criterion used in resident surveys, professional key informants are chosen on the basis of their knowledge about neighbourhood (dis)organisational processes. Key informants are thus field experts. The point of departure is that the privileged witness represents an important additional information to the more established resident surveys. The importance of this principle has already been underscored by Campbell (1955: 340) who stated “if the use of informants as a social science research tool is to be developed, it seems likely that principles of optimal selection will have to be developed”. The principle of optimal selection should ensure that the knowledge of professional key informants exceeds the knowledge of ordinary residents.

The advantage of using key informants is strongly dependent on the quality of the information provided by these individuals. This is especially important when quantitative data are to be provided by the informants (Kendall and Lazarsfeld, 1950). Sudman and Bradburn (1974) have identified an important issue concerning the reliability and validity of the data thus obtained: measurement error increases when informants have only vague knowledge of the topics being explored. A thorough selection is thus necessary when using the technique of KIA. Results are based on the informant’s ability to observe and perceive the underlying social processes. In this study we therefore explicitly posed an initial question to these respondents: do they consider themselves able to answer rating scales on social trust, informal social control and disorder? There is little reason to believe that key informants should be less emotional than local residents about social trust, informal social control and disorder, but in spite of this they should be well informed because of their function. The principle of self-selection is hypothesized to be an important filter question to rule out bias.

Ecological reliability, convergent validity and ecological stability

Measurement can be described as the systematic assignment of numbers to variables to represent features of persons, objects or events (Vandenberg and Lance, 2000). In ecological research one aims at measuring characteristics of ecological settings. Raudenbusch and Sampson (1999) coined the term 'ecometrics' to refer to the art of measuring characteristics of ecological units. Ecological entities involve street-blocks, neighbourhoods, postal code areas or even units at a higher level of analysis. One major issue in ecological research is the question of how close we can get to the measurement of characteristics of ecological units, rather than the measurement of characteristics of respondents answering observational questions on characteristics of ecological entities. After all, individual respondents are providing data to assess characteristics measured at higher levels of aggregation. The perception of individuals determines the ecological survey-based measure. To what extent is it possible to express social processes such as social trust, shared expectations for informal social control and disorder in numbers?

It is clear that a decisive criterion is needed to evaluate the quality of measures of sociological properties of geographical areas. So far, this has been done in the psychometrical tradition by (a) using reliable and valid measures at the respondent's level and by (b) using multilevel modelling to evaluate the ecological reliability of measurement scales created at the individual level (Raudenbush and Sampson, 1999). We highlight the difference between psychometric and ecological reliability. A reliable psychometric scale consists of a set of items that meet the demands of internal consistency. Reliability can be analysed using factor analysis of the observational questionnaire-items and by computing *Cronbach's alpha*, one of the most well-known (conservative) estimators of scale reliability. To assess ecological reliability, Raudenbush and Sampson introduced the *lambda parameter* (Raudenbush and Sampson, 1999). This measure is calculated based on the same components as the *intraclass correlation coefficient* (ICC) and, hence, can be deduced from that parameter. *Lambda* standardizes the variance components for the average sample size within ecological units (average number of individuals per ecological units).

A reliable measure is not necessary a valid measure; accordingly it is necessary to test the ecological validity of area level aggregates. Validity refers to the absence of systematic measurement error. Thus, the measure should measure what it claims to measure. In the present study the focus is on the construct validity of measures (Markus and Lin, 2010). However, the principle of construct validity seems so straightforward that one may easily be misguide by its simplicity. Construct validity is obtained when an ecological measure correlates as highly as would be expected based on theoretical expectations. According to measurement theory, construct validity is obtained when a construct is studied in a network of 'causal arrows' that represent an established theoretical model, and when the relationships between the effect parameters have the expected magnitude and direction (positive or negative). However, in many criminological studies the validity of a measure is often demonstrated by looking at bivariate correlations between constructs. Therefore construct validity is often limited to correlational validity (Meng et al., 1992) or convergent validity.

Lastly, ecological stability will be assessed to gain insight in the stability of social processes in ecological units. Although ecological measures of social processes might change over time, we assume that these measures are relatively stable (at least in the short term) due to its relation with structural factors.

Data and methods

The data collection for this study took place in Ghent, Belgium. The data were collected for the purposes of the Social capital and Well-being In Neighborhoods in Ghent (SWING) study (see Hardyns et al., 2015 for the study protocol). Ghent is the second largest city of Belgium and is located in the southwest of the country with 261,483 residents in 2019 and a surface of 158 km^2 (±1,655 residents/km^2).

Data collection methods

The SWING study consists of multiple successive cross-sectional waves of data collection in neighbourhoods in Ghent. In each cross-sectional wave, multiple methods of data collection were used, however, not every data collection method was used each year. In this study, we use data from the 2012 wave (inhabitant survey and key informant survey), the 2013 wave (inhabitant survey and key informant survey), and the 2014 wave (key informant survey). In total, 1,685 inhabitants (two waves) and 1,485 key informants (three waves) from 91 neighbourhoods were questioned. These neighbourhoods were operationalized as statistical sectors (see Appendix 1 for descriptive statistics), which are comparable to census tracts in the United Kingdom or United States. These units of analysis are the smallest units on which administrative data are available in Belgium (Hardyns et al., 2015).

In 2012 and 2013, data were collected in 91 neighbourhoods (**T1**).[2] The 2012 and 2013 neighbourhood inhabitant surveys yield 1,685 respondents which translates into an average number of respondents per neighbourhood of 18.52. The 2012 and 2013 key informant surveys yield 754 respondents, which is equivalent to an average number of respondent per neighbourhood of 8.29. In 2014, data were collected in the same 91 neighbourhoods (**T2**). The 2014 key informant survey yields 731 respondents, which is an average of 8.03 respondents per neighbourhood. The data collection in two different periods of time (T1 and T2) enables us to assess ecological stability (see below). In Appendix 2, socio-demographic characteristics of the respondents, both inhabitants and key informants, can be found. With T1 and T2 we refer to the two respective full data collection waves spanning the 91 neighbourhoods. T1 refers to the waves of 2012 and 2013. T2 refers to the wave of 2014.

In addition to these survey-based measures of social processes, we collected secondary data for the same 91 neighbourhoods. In particular, we collected (1) data on structural background variables on the community structure and (2) police-registered crime data.

Sampling of neighbourhood inhabitants

The neighbourhood inhabitants in this study were selected based on a randomized sample that was drawn from the municipal registry. This sample was representative for the composition of each neighbourhood (stratified by sex, age and current nationality). The inclusion criteria for neighbourhood inhabitants were: (1) being older than 18 years, (2) not living in an institutional setting (e.g., a home for the elderly, prison), and (3) having sufficient knowledge of the Dutch language to complete the questionnaire. Information on the first two criteria were derived from the municipal registry and taken into account in the sampling, the final criterion (language proficiency) was determined at the moment of first contact. The ambition was reach 20 inhabitants in each neighbourhood, for each wave. The overall response rate of inhabitants in the SWING study was 47.89%.

Sampling of key informants

In contrast to the sampling of neighbourhood inhabitants, key informants are not random sampled. The key informants were purposely chosen by the interviewers based on their supposed knowledge about the social processes that are at stake in this study. The selection of

[2] In the 2012 wave valid data were collected in 50 neighbourhoods. In the 2013 wave valid data were collected in 41 neighbourhoods.

‘good’ key informants was covered during the interviewer training. In addition, each interviewer was provided with a detailed but non-limitative list of possible functions of held by key informants. However, interviewers were encouraged to select other key informants with supposedly good knowledge on the social processes. Our goal was to obtain a heterogeneous set of eight to ten key informants per neighbourhood. The inclusion criteria for key informants were: (1) being older than 18 years, (2) having sufficient knowledge of the Dutch knowledge to complete the questionnaire, and (3) being in a work position that presumes an above average knowledge of the social processes in one of the neighbourhoods studied.

Due to the position of a key informant within a community, their perception on social processes (and, hence, the validity of the subsequent measures) can be affected. By accounting for diversity of the key informants, this effect will be reduced to a minimum. To assess the heterogeneity of the key informants quantitatively, we propose the use of a diversity index. In this study, we use the Simpson diversity index (Simpson, 1949), see equation 1, to assess the average diversity (*D*) of key informants per neighbourhood. Simpson’s diversity index takes into account the number of key informants present, as well as the relative abundance of each category of key informant. The value ranges between 0 and 1, with higher values indicating more diversity.

$$D = 1 - \left(\frac{\sum n(n-1)}{N(N-1)}\right) \tag{1}$$

To calculate the Simpson diversity index, it is necessary to categorise the population into meaningful groups that are capable to express the heterogeneity of the population. We have categorised key informants (before data collection) into ten groups: local organisation/shop (n=572), catering industry (n=188), service sector (n=193), social work (n=135), (para)medical sector (n=223), police and security (n=39), childcare and youth care (n=39), construction industry (n=18), financial sector (n=53), and primary sector (n=22). Only three key informants could not be categorised in one of these groups. For T1, the Simpson diversity index per neighbourhood ranges from 0.25 to 0.96 and the mean Simpson diversity index for all 91 neighbourhoods is 0.77 (SD=0.14). For T2, the Simpson diversity index per neighbourhood ranges from 0.25 to 1.00 and the mean Simpson diversity index for all 91 neighbourhoods is 0.76 (SD=0.14). The mean Simpson diversity indices suggest a diverse and heterogeneous selection of key informants.

Measures

Social trust, informal social control and disorder[3] were questioned in the same manner in both the neighbourhood inhabitant survey and the key informant survey. The question wording, coding and the measure of internal consistency (Cronbach's alpha) of the constructs can be found in Appendix 3. These analyses reveal high scale reliability scores (≥.74).

For the structural background variables on the community structure, we have used data from the city of Ghent from the years of the start of the respective data collection waves (both 2012 and 2014). The unemployment rate is used as a proxy measure of economic deprivation. The percentage non-Belgian inhabitants is used as a proxy measure of immigrant concentration. The choice for these structural background variables is rather pragmatic, but allows for a partial assessment of convergent validity at the census tract level. The police-registered crime data dates from 2013. Crime rates (per 1000 neighbourhood inhabitants) were computed and further analysed using exploratory factor analysis. The measure of 'vandalism' is a sum score consisting of several crime types: vandalism against vehicles, vandalism against buildings, use and possession of drugs, noise pollution and alcohol abuse. An exploratory factor analysis revealed that all items load sufficiently high on one factor (factor loadings ≥0.54). The measure of 'violent crime' is a sum score consisting of two crime types: assault and battery, and threat. Based on a factor analysis, these items all load on the same factor (factor loadings ≥0.95).

Analytical strategy

SPSS Statistics (version 25) is used to conduct the analyses. To assess the **scale reliability** or internal consistency for the scale constructs, *Cronbach's alpha* was computed. The **ecological reliability** of neighbourhood processes was assessed by means of the *intraclass correlation coefficient* (ICC) and the *lambda parameter* (Raudenbush and Sampson, 1999). The ICC is represented in equation 2, where σ^2 represents the variance and B and W respectively stand for between and within groups (Heck et al., 2010).

$$\rho = \sigma_B^2/(\sigma_B^2 + \sigma_W^2) \tag{2}$$

[3] Mention that physical disorder is only measured in the key informant survey and not in the neighbourhood inhabitant survey.

The *lambda parameter* standardized the variance components computed by means of the ICC for the average number of individuals within a neighbourhood (Heck et al., 2010; Snijder and Bosker, 2012):

$$\lambda = \frac{\sigma^2}{\sigma_B^2 + (\sigma_W^2/n_j)} \tag{3}$$

This *lambda parameter* for ecological reliability can also be deduced from the ICC statistic (Snijder and Bosker, 2012: 26):

$$\lambda = \frac{n_j\rho}{1 + (n_j - 1)\rho} \tag{4}$$

The lambda parameter has, analogous to its psychometric equivalent Cronbach's alpha, a value between 0 and 1. A lambda value of 0.80 or more indicates a good to excellent ecological reliability, but as this measure is equivalent to the Cronbach's alpha, a value of 0.70 or higher ensures an acceptable level of ecological reliability.

The ICC can be determined from a two-level hierarchical linear regression intercept-only model (i.e., a model with no covariates). We used a two-level multilevel regression model with neighbourhoods as level-2 units and individuals (either inhabitants or key informants) as level-1 units. We used restricted maximum likelihood (REML) as the estimation method (Heck et al., 2010). These multi-level analyses account for the nested data structure of people within neighbourhoods. For the data based on the neighbourhood inhabitant survey, we applied an ecometrics approach. Specifically, we used the level-2 predicted values for the outcome variables, controlling for individual level covariates (the characteristics used for the stratified random sampling of neighbourhood inhabitants like gender, age and nationality at birth). This is an established way to control for neighbourhood compositional effects. Interestingly, correcting for compositional effects did not lead to different scores of the neighbourhood level measures, as the predicted values correlate almost perfect with the predicted values from the conditional model (between 0.97 and 0.98 on the individual level, and 0.99 or 1.00 at the neighbourhood level, all significant on the 0.001-level). Finally, to assess **convergent validity** and **ecological stability** of ecological processes we used correlational statistics.

Results

Ecological reliability

Table 1 contains the results of the computation of the parameters of ecological reliability: the intraclass correlation coefficient (ICC) and the lambda parameter. A closer look at both parameters reveals that there are notable differences between the different methods. As can be observed, the resulting lambda parameters for the measures from the key informant surveys are inherently lower than those from the neighbourhood inhabitant survey. This is caused by the fact that this parameter is standardized by the average number of respondents per ecological unit and, as mentioned above, the average number of key informants is lower than the average number of neighbourhood inhabitants.

First, the measure of social trust is relatively stable across the different methods. As can be seen in Table 1, between 13 and 15 percent of the variance of social trust is due to differences between neighbourhoods. Second, the ICC of informal social control is comparable over the different methods when comparing T1 and T2 (between 10 and 11 percent). However, the ICC value of informal social control is clearly lower (around 5 percent) on T1 when measured by key informants. For both social trust and informal social control, the lambda parameter indicates that these constructs cannot be measured reliably at the neighbourhood level. Third, disorder is measured remarkably more ecological reliable than the aforementioned organisational processes (lambda values between 76 and 92 percent), and between 28 and 38 percent of the variance is due to differences between neighbourhoods.

	Social trust	Informal social control	Social disorder	Physical disorder
Inhabitant survey T1				
ICC	0,13	0,11	0,37	-
lambda	0,74	0,71	0,92	-
Key informant survey T1				
ICC	0,15	0,05	0,32	0,28
lambda	0,59	0,29	0,80	0,76
Key informant survey T2				
ICC	0,13	0,10	0,38	0,35
lambda	0,54	0,46	0,83	0,81

Table 1. Reliability of ecological constructs (intraclass correlation coefficient and lambda parameter).

In Figure 1 the sensitivity of the lambda parameter is presented as a function of the average number of respondents per ecological unit. The finding regarding the ecological reliability of the measures of social trust and social disorder are similar to the findings of Raudenbush and Sampson (1999). While utterly small differences remain, the reliability of the measures of social trust and social disorder is highly similar, regardless of the method of data collection (neighbourhood inhabitant survey versus key informant survey). The results are substantially different with regard to our measures of informal social control. To obtain ecological reliable measures much more key informants are necessary (between 22 and 46 key informants to obtain a lambda value of 0.70).

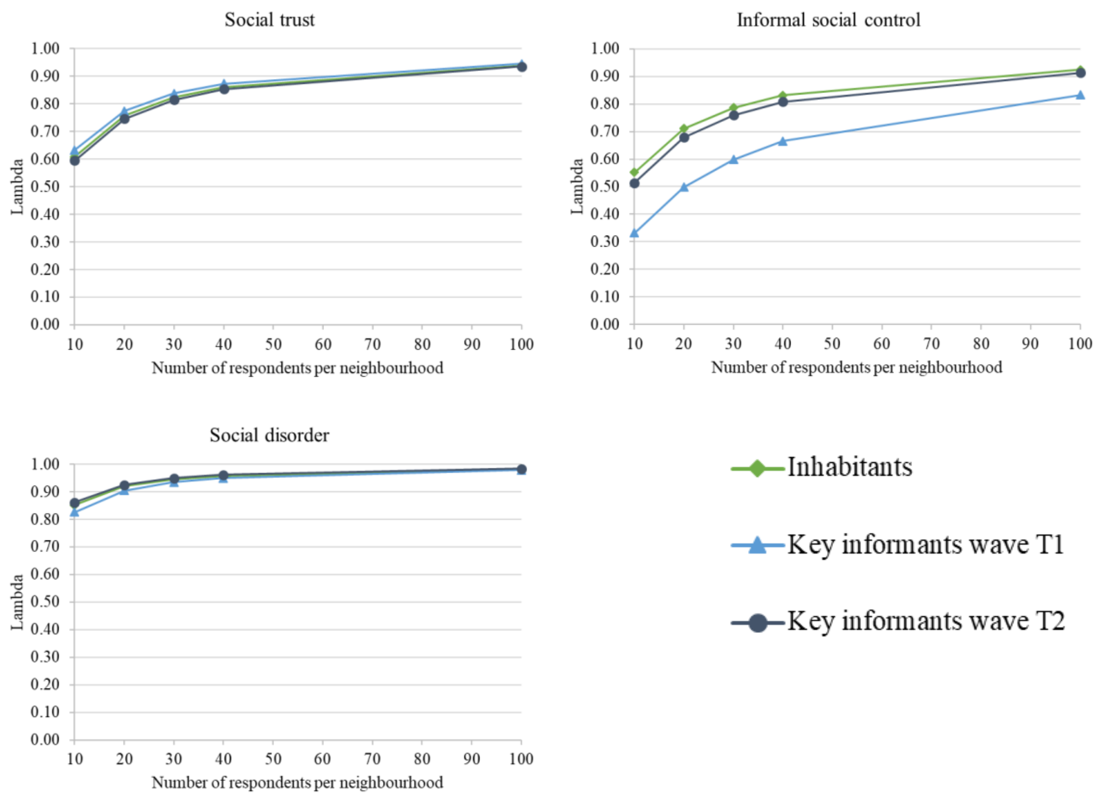

Figure 1. Sensitivity of measures of ecological processes as a function of the number of respondents.

Convergent validity

As mentioned above, convergent validity is obtained when an ecological measure correlates as high as would be expected based on theoretical expectations. In this study, the convergent validity is assessed in two manners. First, the associations between the measures of the ecological processes within the same methods are assessed. Second, the association between measures of the ecological processes between the different methods are assessed. Third, the associations with both structural ecological determinants and potential negative consequences of neighbourhood (dis)organisational processes, from other data sources, are used to test convergent validity.

Table 2 contains a correlation matrix of the ecological processes within the same data collection methods. The correlation between both dimensions of the collective efficacy concept is noteworthy. Based on the neighbourhood inhabitant survey, a strong, significant and positive correlation was found between social trust and informal social control. This result is comparable to the correlation of 0.80 (p<0.001) that was originally reported by Sampson and colleagues (1997). Based on the key informant analysis, only a substantially small relationship was found between social trust and informal social control. It is found that measures of the informal control dimension of collective efficacy perform very badly. Informal social control as measured by the key informant survey is not correlated with the other characteristics (in terms of magnitude) as one would expect based on the ecological model of collective efficacy. The correlation between social trust and disorder (*r* between -.45 and -.58), is in line with the theoretical framework.

			Social trust	Informal social control	Social disorder	Physical disorder
Inhabitants	T1	Social trust		0.76***	-0.67***	
		Informal social control			-0.63***	
		Social disorder				
Key informants	T1 & T2	Social trust		0.23*	-0.58***	-0.56***
		Informal social control	0.21 n.s.		-0.11 n.s.	-0.26*
		Social disorder	-0.58***	-0.20 n.s.		0.78***
		Physical disorder	-0.45***	-0.05 n.s.	0.83***	

Table 2. Correlation matrix ecological processes (Note. For key informants: T1 is above the diagonal, T2 is below the diagonal).

Second, we examined to what extent similar ecological processes measured by means of different data collection methods are associated with each other. The results show that both neighbourhood level social trust and social disorder are strongly and positively associated. However, informal social control is correlated weakly with the other constructs across the different methods (see Table 3).

		Neighbourhood inhabitants T1		
		Social trust	Informal social control	Social disorder
Key informants T1	Social trust	0.46***		
	Informal social control		0.22*	
	Social disorder			0.74***
Key informants T2	Social trust	0.54***		
	Informal social control		0.21*	
	Social disorder			0.76***

Table 3. Correlational validity ecological processes neighbourhood inhabitants and key informants (T1 and T2).

Third, we assessed to what extent these (dis)organisational processes measured by means of different methods are associated with both structural neighbourhood characteristics and police-registered crime data (see Table 4). It was found that these parameters all correlated well and as expected, except for informal social control measured by means of a key informant survey. This corroborates the fact that the concept of informal social control is not validly measured in this manner.

The overall picture that emerges is that the correlations between on the one hand the neighbourhood processes and on the other hand the structural determinants and negative (crime-related) outcomes based on the key informant survey are not as strong as measured by neighbourhood inhabitant surveys. However, the effects are still moderate to strong in general.

	Social trust	Informal social control	Social disorder	Physical disorder
Inhabitant survey T1				
Community structure				
Unemployment rate	-0.66***	-0.65***	0.82***	
% non-Belgian	-0.56***	-0.58***	0.82***	
Crime variables				
Vandalism	-0.30**	-0.32**	0.51***	
Violent crime	-0.31**	-0.39***	0.46***	
Key informant survey T1				
Community structure				
Unemployment rate	-0.54***	-0.23*	0.75***	0.62***
% non-Belgian	-0.49***	-0.23*	0.66***	0.61***
Crime variables				
Vandalism	-0.33***	-0.10 n.s.	0.26*	0.24*
Violent crime	-0.22*	-0.10 n.s.	0.25*	0.25*
Key informant survey T2				
Community structure				
Unemployment rate	-0.58***	-0.22*	0.70***	0.62***
% non-Belgian	-0.51***	-0.13 n.s.	0.66***	0.63***
Crime variables				
Vandalism	-0.36***	-0.10 n.s.	0.37***	0.31**
Violent crime	-0.44***	-0.06 n.s.	0.33***	0.23*

Table 4. Correlation matrix with structural ecological determinants (community structure) and consequences (crime variables).

Ecological stability across time

Table 5 contains a correlation matrix with the measures of the key informant surveys on T1 and T2. This analysis confirms the previous finding that informal social control is neither measured valid, nor reliable. This is remarkable, because, although one could argue that social processes are subject to changes in the ecosystem due to various social, societal and environmental changes (Hardyns et al., 2018), these could not be so volatile that they change drastically from one year to another, since these changes are related to structural determinants. When examining the correlations between social trust and both measures of disorder, we come to the conclusion that a strong and positive correlation exists. This is highly suggestive for the relative stability of these processes on the neighbourhood level.

		Key informants T1			
		Social trust	Informal social control	Social disorder	Physical disorder
Key informants T2	Social trust	0.58***			
	Informal social control		0.01 n.s.		
	Social disorder			0.78***	
	Physical disorder				0.59***

Table 5. Correlation matrix stability ecological processes measured with key informant technique.

Conclusion and discussion

The results of this study clearly suggest that the design of ecologically reliable measures is not only affected by choices made at the item level. Additionally, constructing ecologically reliable measures of neighbourhood processes is also affected by the eye of the beholder. This is consistent with measurement error theory (e.g. Fuller, 1987) and survey response models (Tourangeau, Rips and Rasinski, 2000; Zaller and Feldman, 1992). However, how straightforward this idea may be, it is important to get insight in the magnitude of measurement problems when applying a method of data collection. Some theoretical constructs are more suitable to measure with the KIA technique than others. This study showed that the KIA technique is well-suited to measure the presence of negative outcomes (e.g., social and physical disorder). In the present study, surveying key informants created ecological reliable, valid and stable measures of disorder. The measurement of social trust is also possible by both probing key informants and neighbourhood inhabitants, however, there are still a considerable number of key informants needed to obtain ecological reliability. Informal social control on the other hand is neither reliable nor validly measured by means of surveying key informants.

Prior research already pointed to the observation that the measurement of collective efficacy depends on the specific cultural context and cannot simply be transferred and applied to other countries and contexts (Hardyns, Snaphaan, Pauwels, et al., 2019; Reisig and Cancino, 2004; Zhang et al., 2007, 2009). The observed differences in the outcomes of collective efficacy research may therefore be partially explained due to different measurements of the concept or the specific contexts (e.g., metropolitan cities versus regional cities, US versus Western Europe). However, we cannot ignore the argument that key informants may not be suited to measure informal social control. After all, (neighbourhood) informal social control is probed by asking the respondents how likely it is it that they could count on the neighbours for intervening when problems of crime and disorder arise in the neighbourhood. As we demonstrated that key informants have specific socio-demographic characteristics, other than the average inhabitant (see Appendix 2), this difference may affect the perception on informal social control (i.e. scorer disagreement).

Apart from the reliability and validity of the measurements by means of the KIA technique, this method has significant advantages on its own. The problem of shared survey-method variance (Sampson and Raudenbush, 1999; Taylor, 1999) is eliminated by using different methods in composing the scales used. In addition, the condition of stratified random sampling does – compared to an inhabitant survey – not have to be met and therefore a lower number of respondents can be sampled in order to obtain sufficient reliable ecological measures. For example, in measuring social disorder a number of ten respondents per neighbourhood would be sufficient (lambda parameter ≥.80). However, a stratified random sampling design cannot be applied in an appropriate manner to such a low sample size per unit. There are simply to many inhabitants needed to obtain a representative sample of the unit of analysis that is studied. Hence, the method of KIA is more cost effective when studying disorder.

Although we believe that this contribution yield informative results to enhance the methodological toolkit of social scientists, we must address the limitations of the present study. For the purpose of this inquiry, we pooled data of 2012 and 2013 to one dataset (T1). In order to do so, we had to accept the auxiliary hypothesis that remarkable changes in the social climate of neighbourhoods did not take place between 2012 and 2013, but preferably all data were collected in one wave. Another problem arises when survey data is used. In spite of the use of the 'ecometrics approach' to capture neighbourhood processes, individual perceptions still lie at the core of the neighbourhood measures. We need to take into account that this creates more bias than unobtrusive measures. However, truly unobtrusive measures free of measurement error do not exist, neither in the physical sciences as in the social sciences. Therefore, increasing insights in the nature of measurement issues remain paramount. Research on neighbourhood processes and mechanisms, just like research on individual processes and mechanisms, needs calibrated instruments. Additionally, while we were able to use multiple methods at several points in time, we acknowledge the restriction that our results are restricted to one urban context (Ghent) and one unit of analysis (neighbourhoods).[4] However, like multiple independent variables are included in statistical analysis to assess statistical significance, scholars should consider multiple units of analysis to assess to what extent the choice for a particular ecological units biases the outcomes (Hardyns, Snaphaan and Pauwels, 2019; Rengert and Lockwood, 2009). Noteworthy is that prior research indicates that KIA is also applicable in micro places (Gerell, 2015, 2017).

Rather a general, but important methodological, limitation is that the observed differences between the measures at T1 and T2 might be due to other changes than a "real" change of the measures of ecological processes. First, differences might be due to changes in the factor structure (i.e. item parameter drift) at the individual level of key informants; just because it are other key informants in T1 and T2. Second, changes in the ecological reliability might be due to a shift of meaning of the concept at hand (e.g., changes of the level of social trust or informal social control in certain neighbourhoods). However, this points to the importance of (the replication of) inquiries into measurement error at different points in time. Studying measurement issues should not be seen as a goal on its own, but as a necessary practice to identify the least problematic available data and methods, i.e. containing as less measurement errors as possible, to represent concepts in ecological research. The empirical testing of substantial theories is strongly dependent of our current state-of-the-art knowledge of best practices in measuring social and ecological processes.

From a measurement error perspective, there are two separate important avenues to improve (our knowledge on) the measurement of ecological processes and hence our theories and insights in the area of ecological research. The first one is related to improving conventional

---

[4] Thus, the Modifiable Areal Unit Problem (MAUP) is left unaddressed in this study.

data collection methods and measurement instruments, the second one is related to the use of new and emerging methods and instruments.

First, related to the conventional data collection methods and measurement instruments, there are other important methodological questions that are not addressed in this study. It is important to consider the question of the 'right' ecological level. As Groff (2018: 113) puts it: "social processes exist on a continuum of concreteness from the individual to the micro level of behaviour settings to neighborhoods. These levels do not operate in isolation; rather there is significant interaction among them". Thus, scholars should apply the best available ('gold standard') methodological approaches that best fit their theoretical perspective. It is important to acknowledge that the processes that produce measurement errors are not mysterious, but produced by real psychological processes. Hence, these processes can be studied, understood, and reduced to a minimum (Schmidt and Hunter, 1999). In this study, scorer disagreement is left unaddressed, but might have repercussions for the results. Future research should therefore address these important questions. We also hypothesized in this study that the principle of self-selection was an important filter to rule out bias, but we did not assess this empirically. Further, two recommendations need to be made. As initiated in this study, we recommend the use of a diversity index (e.g. Simpson diversity index) to quantify the (average) diversity of selected key informants per ecological unit. Last, we should acknowledge that the KIA technique is only one tool in the toolkit of the social scientist. Nevertheless, this technique can potentially also be used to measure other ecological processes than measured in this study.

Second, new and emerging data sources, and innovative methods of data collection, -processing and -analysis methods provide (underexplored) opportunities in measuring (dis)organisational processes in an innovative manner. The use of these new and emerging data sources is proliferating within spatio-temporal approaches of criminology (Snaphaan and Hardyns, 2019). For example, the analysis of secondary street-level imagery, like Google Street View data, have proven to be a valid alternative for in-situ observations (Marco et al., 2017; Odgers et al., 2012). The domains of machine learning and computer vision enable scholars to process and analyse these types of data automatically (Mittal et al., 2016; Sukel et al., 2019). New technologies also allow scholars to gather primary imagery data at a much larger scale. However, this line of inquiry poses other challenges from a legal, technological and safety perspective (Grubesic et al., 2018). Besides the methodological advantages, these new data sources are also promising from a conceptual perspective, as these allow to distinct between both perceptual and non-perceptual measures of disorder (cfr. Sampson and Raudenbush, 1999) at an unprecedented scale. The key disadvantage of these new and emerging data sources, as with observations in general, is that these methods cannot validly capture the rich theoretical concepts that require neighbourhood residents' perspectives (Raudenbush and Sampson, 1999). As Raudenbush and Samson (1999: 11) mention: "If researchers rely entirely on observations, there is a danger that they will misinterpret the significance of observable conditions such as physical disorder, building conditions, and land use". It is thus possible to measure (dis)organisational processes in an innovative manner, but the question remains to what extent we can validly measure social processes with these new and emerging data sources and methods. It may be clear that here lies a large unexplored area, that yields interesting avenues for future research.

References

Bruinsma GJN and Johnson SJ (2018) Environmental criminology: History, scope and state of the art. In Bruinsma GJN and Johnson SJ (eds) *The Oxford Handbook on Environmental Criminology*. New York: Oxford University Press, pp. 1-31.

Brunton-Smith I, Sturgis P and Leckie G (2018) How collective is collective efficacy? The importance of consensus in judgments about community cohesion and willingness to intervene. *Criminology* 56(3): 608-637.

Bunge M (2003) *Emergence and convergence: Qualitative novelty and the unity of knowledge*. Toronto: University of Toronto Press.

Bursik RJ and Grasmick HG (1993) *Neighborhoods and crime*. New York: Lexington.

Campbell DT (1955) The informant in quantitative research. *American Journal of Sociology* 60(4): 339-342.

Campbell DT and Fiske DW (1959) Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin* 56(2): 81-105.

Earls FJ, Raudenbush SW, Reiss AJ and Sampson RJ (2005) *Project on Human Development in Chicago Neighborhoods (PHDCN): Systematic Social Observation, 1995. ICPSR13578-v1*. *Documentation.* Ann Arbor, MI: Inter-university Consortium for Political and Social Research. Available at: http://doi.org/10.3886/ICPSR13578.v1

Fuller WA (1987) *Measurement error models*. New York, NY: Wiley.

Gerell M (2017) *Neighborhoods without community: Collective efficacy and crime in Malmö, Sweden.* Doctoral dissertation, Malmö University, Sweden. Available at: http://muep.mau.se/bitstream/handle/2043/15962/systematicsocialobservationofphysicaldisorderininner-cityurbanneighborhoodsthroughgooglestreetview.pdf?sequence=2

Gerell, M (2015) Collective efficacy, neighborhood and geographical units of analysis: Findings from a case study of Swedish residential neighborhoods. *European Journal on Criminal Policy and Research* 21(3): 385-406.

Groff ER (2018) Do we really need collective social process to understand why crime occurs and offenders commit crime? In Bruinsma GJN and Johnson SD (eds) *The Oxford Handbook of Environmental Criminology*. New York: Oxford University Press, pp. 105-118.

Grubesic TH, Wallace D, Chamberlain AW and Nelson JR (2018) Using unmanned aerial systems (UAS) for remotely sensing physical disorder in neighborhoods. *Landscape and Urban Planning* 169: 148-159.

Hardyns W, Pauwels LJR and Heylen B (2018) Within-individual change in social support, perceived collective efficacy, perceived disorder and fear of crime: Results from a two-wave panel study. *British Journal of Criminology* 58(5): 1254-1270.

Hardyns W, Snaphaan T, Pauwels LJR, Vyncke V and Willems S (2019). A multilevel analysis of collective efficacy, neighborhood disorder, and individual social capital on avoidance behavior. *Crime & Delinquency* 65(7): 994-1021. doi: 10.1177/0011128718788042

Hardyns W, Snaphaan, T and Pauwels LJR (2019) Crime concentrations and micro places: An empirical test of the "law of crime concentration at places" in Belgium. *Australian & New Zealand Journal of Criminology* 52(3): 390-410. doi: 10.1177/0004865818807243

Hardyns W, Vyncke V, Pauwels L and Willems S (2015). Study protocol: SWING – social capital and well-being in neighborhoods in Ghent. *International Journal for Equity in Health,* 14: 1-10. doi:10.1186/s12939-015-0163-1

Heck RH, Thomas SL and Tabata L (2010) *Multilevel and longitudinal modeling with IBM SPSS*. New York: Routledge.

Hipp JR (2016) Collective efficacy: How is it conceptualized, how is it measured, and does it really matter for understanding perceived neighborhood crime and disorder? *Journal of Criminal Justice* 46: 32-44.

Hoeben EM, Steenbeek W and Pauwels LJR (2018). Measuring disorder: Observer bias in systematic social observations at streets and neighborhoods. *Journal of Quantitative Criminology* 34(1): 221-249.

Kendall PL and Lazarsfeld PF (1950). Problems of survey analysis. In Merton RK and Lazarsfeld PF (eds) *Continuities in Social Research*. Glencoe, IL: The Free Press, pp. 133-196.

Kronkvist K (2013) *Systematic social observation of physical in inner-city urban neighborhoods through Google Street View: The correlation between virtually observed physical disorder, self-reported disorder and victimization of property crimes.* Master thesis, Malmö University, Sweden. Available at: http://muep.mau.se/bitstream/handle/2043/15962/systematicsocialobservationofphysicaldisorderininner-cityurbanneighborhoodsthroughgooglestreetview.pdf?sequence=2

Kubrin CE and Wo JC (2015) Social disorganization theory's greatest challenge: Linking structural characteristics to crime in socially disorganized communities. In Piquero AR (ed), *The Handbook of Criminological Theory*. Hoboken, NJ: John Wiley & Sons, pp. 121–136.

Kurtz LR (1984) *Evaluating Chicago sociology: A guide to the literature, with an annotated bibliography*. Chicago: University of Chicago Press.

Larson R and Csikszentmihalyi M (1983) The experience sampling method. *New Directions forMethodology of Social & Behavioral Science* 15: 41-56.

Marco M. Gracia E, Tomás JM and López-Quílez A. (2015) Assessing neighborhood disorder: Validation of a three-factor observational scale. *The European Journal of Psychology Applied to Legal Context* 7(2): 81-89.

Markus KA and Lin C (2010) Construct validity. In Salkind NJ (ed) *Encyclopedia of research design.* Thousand Oaks, CA: SAGE Publications, pp. 229-233.

Mastrofski SD, Parks RB and McCluskey JD (2010) Systematic social observation in criminology. In Piquero AR and Weisburd D (eds), *Handbook of quantitative criminology*, Cham: Springer, pp. 225-247.

Mazerolle L, Wickes R and McBroom J (2010) Community variations in violence: The role of social ties and collective efficacy in comparative context. *Journal of Research in Crime and Delinquency* 47(1): 3-30.

Meng X-L, Rosenthal R and Rubin DB (1992) Comparing correlated correlation coefficients, *Psychological Bulletin*, 111(1): 172-175.

Mittal G, Yagnik KB, Garg M and Krishnan NC (2016) SpotGarbage: Smartphone app to detect garbage using deep learning. In: *UbiComp Conference*, Heidelberg, Germany, pp. 940-945.

Oberwittler D and Wikström P-OH (2009) Why smaller is better: Advancing the study of the role of behavioral contexts in crime causation. In Weisburd D, Bernasco, W and Bruinsma GJN (eds) *Putting Crime in Its Place: Units of Analysis in Geographic Criminology*. New York: Springer Science Business Media, pp. 36-60.

Odgers CL, Caspi A, Bates CJ, Sampson RJ and Moffitt TE (2012) Systematic social observation of children's neighborhoods using Google Street View: A reliable and cost-effective method. *Journal of Child Psychology and Psychiatry* 53(10): 1009-1017. https://doi.org/10.1111/j.1469-7610.2012.02565.x

Park RE, Burgess EW and McKenzie RD (1967 [1925]) *The city: Suggestions for investigation of human behavior in the urban environment*. Chicago: University of Chicago Press.

Park, RE and Burgess, EW (1929) *Introduction to the science of sociology*. Chicago: University of Chicago Press.

Pauwels L and Hardyns W (2009) Measuring community (dis)organisational processes through key informant analysis. *European Journal of Criminology* 6(5): 401-417. doi: 10.1177/1477370809337878

Raudenbush SW and Sampson RJ (1999) Ecometrics: Toward a science of assessing ecological settings, with application to the systematic social observation of neighborhoods. *Sociological Methodology* 29(1): 1-41.

Reisig MD and Cancino JM (2004) Incivilities in nonmetropolitan communities: The effects of structural constraints, social conditions, and crime. *Journal of Criminal Justice* 32: 15-29.

Rengert GF and Lockwood B (2009). Geographical units of analysis and the analysis of crime. In Weisburd D, Bernasco, W and Bruinsma GJN (eds) *Putting crime in its place: Units of analysis in geographic criminology*. New York: Springer Science Business Media, pp. 109-122.

Sampson RJ (2012) *Great American city: Chicago and the enduring neighborhood effect*. Chicago: University of Chicago Press.

Sampson RJ and Groves WB (1989) Community structure and crime: Testing social-disorganization theory. *American Journal of Sociology* 94(4): 774-802.

Sampson RJ and Raudenbush SW (1999) Systematic social observation of public spaces: A new look at disorder in urban neighborhoods. *American Journal of Sociology* 105(3): 603-651.

Sampson RJ, Raudenbush SW and Earls F (1997) Neighborhoods and violent crime: A multilevel study of collective efficacy. *Science* 277(5328): 918-924.

Schmidt FL and Hunter JE (1999) Theory testing and measurement error. *Intelligence* 27(3): 183-198.

Shaw CR and McKay HD (1942) *Juvenile Delinquency and Urban Areas.* Chicago: University of Chicago Press.

Snaphaan T and Hardyns W (2019) Environmental criminology in the big data era. *European Journal of Criminology.* doi: 10.1177/1477370819877753

Simpson, EH (1949) Measurement of diversity. *Nature* 163: 688. https://doi.org/10.1038/163688a0

Snijders TAB and Bosker RJ (2012) *Multilevel analysis: An introduction to basic and advanced multilevel modeling*. Los Angeles: SAGE.

Steenbeek W (2011) *Social and physical disorder. How community, business presence and entrepreneurs influence disorder in Dutch neighborhoods*. Doctoral dissertation, Utrecht University, the Netherlands.

Sudman S and Bradburn NM (1974) *Response effects in surveys: A review and synthesis.* Chicago: Aldine Publishing Company.

Sukel M, Rudinac S and Worring M (2019) Multimodal classification of urban micro-events. In: *27th ACM International Conference on Multimedia*, Nice, France, pp. 1455-1463. https://doi.org/10.1145/3343031.3350967

Taylor RB (1999) The incivilities thesis: Theory, measurement, and policy. In Langworthy RH (ed), *Measuring what matters.* Washington, D.C.: National Institute of Justice, pp. 65-88.

Thorndike EL (1920) A constant error in psychological ratings. *Journal of Applied Psychology* 4(1): 25-29.

Tiby E and Olsson M (1997). *Lokala problembilder – Nyckelinformanters syn på brott, sociala problem och åtgärder i åtta förorter till Stockholm.* Licentiatsuppsats, Stockholms universitet, Sweden.

Tourangeau R, Rips LJ and Rasinski K (2000) *The psychology of survey response*. Cambridge: Cambridge University Press.

Vandenberg RJ and Lance CE (2000) A Review and synthesis of the measurement invariance literature: Suggestions, practices, and recommendations for organisational research. *Organisational Research Methods* 3(1): 4-70.

Wikström P-O (2007) The social ecology of crime. The role of the environment in crime causation. In Schneider HJ (ed), *Internationales Handbuch der Kriminologie*. Berlin: Walter de Gruyter, pp. 333-357.

Wikström PO and Treiber K (2016) Situational theory. The importance of interactions and action mechanisms. In Piquero A (ed) *The handbook of criminological theory*. Chicester: Wiley Blackwell, pp. 415-443.

Zaller, J and Feldman S (1992) A simple theory of the survey response: Answering questions versus revealing preferences. *American Journal of Political Science* 36(3): 579-616.

Zhang LN, Messner SF and Liu JH (2007) A multilevel analysis of the risk of household burglary in the city of Tianjin, China. *British Journal of Criminology* 47: 918-937.

Zhang LN, Messner SF, Liu JH and Zhuo YA (2009) Guanxi and fear of crime in contemporary urban China. *British Journal of Criminology* 49: 472-490.