# Chapter 2
# Why Small Is Better: Advancing the Study of the Role of Behavioral Contexts in Crime Causation

**Dietrich Oberwittler and Per-Olof H. Wikström**

**Abstract** In this chapter we argue, both from a theoretical (Situational Action Theory) and methodological (homogeneity of environmental conditions) point of view, that small environmental units are preferable to large in the study of environmental effects on crime.

Most empirical research in the field of communities and crime utilizes fairly large spatial units of several thousand residents, such as U.S. census tracts or even clusters of census tracts, thus evoking doubts about internal homogeneity. If geographical areas are heterogeneous in their environmental conditions, associations between structural conditions, social organization, and outcomes such as crime may be clouded or rendered insignificant. On the other hand, due to common financial restrictions, choosing more units often (but not necessarily) imply fewer subjects per units which may cause a 'small number problem', that is, that the prediction of events as rare as crime will lose precision (compared to the use of larger units with more subjects). The question then is how small can you go before this potential problem outweighs the benefits of more homogeneous areas? This chapter assesses the added value of using very small area units in a community survey on environmental influences on crime. This survey was carried out in 2005 as part of the Peterborough Adolescent and Young Adult Development Study (PADS+) and covers the UK city of Peterborough and some rural surroundings. For the purpose of this study, we used the smallest administrative unit which subdivides the city, isolating 550 areas with about 300 residents each. We sampled an average of 13 respondents per unit for a total sample of 6,600 respondents. Multilevel analyses and Raudenbush and Sampson's (1999) ecometric approach are applied to compare the aggregate-level reliability of survey scales on this very small geographical level to the larger spatial level conventionally used for geographical analysis. The results show a considerable increase in between-neighborhood variance, reflecting a higher degree of homogeneity and statistical power for detecting particularly moderate to weak area-level effects. We use the collective efficacy scale and its subscales to illustrate these results.

D. Oberwittler
Department of Criminology, Max Planck Institute for Foreign and International Criminal Law, Freiburg, Germany; University of Freiburg, Germany
e-mail: d.oberwittler@mpicc.de

D. Weisburd et al. (eds.), *Putting Crime in its Place*,

## Introduction

The role of the social environment is probably one of the least understood aspects of crime causation. One important reason for that is the lack of well-developed theoretical models of *how* social environments influence individual acts of crime and the development of crime propensity. Another important reason is the lack of well-developed methodologies to study and measure the influence of social environments on individual acts of crime and the development of crime propensity (Wikström 2007a).

A major aim of *the Peterborough Adolescent and Young Adult Development Study* (PADS+) is to advance theory and methodology in the study of crime causation, with a particular focus on the role of the social environment. A specific aim of this chapter is to advocate, on theoretical and methodological grounds, the advantages of using small area units to study the effects of the social environment on human development and action.

PADS+ is based on a newly developed theoretical framework, the *Situational Action Theory* (see Wikström 2004, 2005, 2006, 2007a, 2007b, 2007c; Wikström and Treiber 2007, 2008) which is specifically designed to address the role of the interaction between individuals and the social environment in crime causation. The cornerstone of the theory is that individual actions (like acts of crime) are an outcome of how people *perceive their action alternatives* and (on that basis) *make choices* when confronted with the particularities of *behavior-settings*. A *behavior-setting* may be defined as the part of the environment which an individual, at a particular moment in time, can access with his or her senses, including any media present (Wikström 2006).

According to the theory, only those individual and environmental characteristics that (directly or indirectly) influence how people perceive action alternatives and make choices are relevant to the study of crime causation. That is, those individual characteristics and experiences, and features of the environments to which an individual is exposed, which (directly or indirectly) influence whether or not he or she sees acts of crime as action alternatives and his or her choice to act upon such perceptions are relevant factors in a theory of crime causation.

In brief, the theory proposes that crimes are moral actions (actions which are guided by rules about what it is right or wrong to do) and have to be explained as such. It further proposes that individual differences in morality (moral values and emotions) and ability to exercise self-control, on the one hand, and the moral contexts (applicable moral rules and their enforcement and sanctions) of the behavior-settings in which individuals take part, on the other, interact to determine what action alternatives an individual will perceive, what choices he or she will make and, consequently, what kind of action will follow (e.g., an act of crime). Individual differences in crime involvement are thus, according to the theory, an outcome of differences in individuals' characteristics and experiences (morality and ability to exercise self-control) *and* their exposure to particular behavior-settings (their moral contexts). In other words, the theory proposes that some kinds of individuals in some kinds of settings are more likely to commit crimes than others, depending on the interaction of their morality (and ability to exercise self-control) and the moral contexts to which they are exposed. Individuals' exposure to behavior-settings is, according to the theory, not only directly relevant to their moral actions but also (in the longer term) to the development of individual characteristics related to their crime propensity (their morality and ability to exercise self-control).

This implies that individuals' exposure to behavior-settings, and changes over time in their exposure to behavior-settings, are key in understanding their crime involvement, and changes over time in their crime involvement (i.e., individual crime trajectories) because they affect the development and change of individuals' crime propensities and exposure to environmental inducements and, particularly, the interaction between the two.

The Situational Action Theory's strong emphasis on the role of the social environment (and its changes) in individuals' involvement, and changing involvement, in crime, highlights the need to adequately measure relevant aspects of the social environment, and its changes. Therefore a specific aim of PADS+ is to utilize better techniques (i) to measure relevant *characteristics* of the social environment (behavior-settings) and (ii) to measure individuals' *exposure to* different social environments (their activity fields) than have been previously used in longitudinal studies of individuals' crime involvement and its development and changes.

Since the theory proposes that human development and action is influenced by the behavior-settings in which individuals take part, it is crucial to develop methods to measure units that approximate behavior-settings. Because behavior-settings are the parts of the social environment the individual can access with his or her senses, these units have to be geographically small. Since the theory also proposes that it is the moral context of behavior-settings that influences an individual's moral action, it is important to develop measures that tap into the moral context of the studied units. In other words, the theory implies that we need to develop measurements which capture the moral context of small units which approximate behavior-settings to adequately study the role of the environment in crime causation.

The theory also proposes that it is the behavior-settings to which an individual actually is *exposed* that are of relevance to his or her development and actions. An individual does not, for example, only act and develop in the area surrounding his or her place of residence (i.e., his or her neighborhood). Thus we also need to develop techniques to measure individuals' exposure to different behavior-settings (within or outside their neighborhoods). The configuration of settings to which an individual is exposed within a certain period of time may be referred to as his or her *activity field* for that period (see further, Wikström 2005, 2006). To measure activity fields we have developed a *space-time budget* technique which measures an individual's hourly exposure, over a specific time period, to different kinds of behavior-settings, i.e., his or her activity field (Wikström and Ceccato 2004; Wikström and Butterworth 2006).

In this chapter, we will deal only with the problems related to advantages and disadvantages of using small units of analysis in ecologically oriented research. The equally important problem of measuring individuals' *exposure* to behavior-settings (through the space-time budget technique) will be dealt with elsewhere.

## The 'Unit of Analysis' Problem

The unit of analysis is rarely a problem when we study individuals. An *individual* is defined by his or her body and when we study individuals we study characteristics of their bodies (e.g., weight and height) or internal to their bodies (e.g., values and emotions). Deciding on the best unit of analysis is also relatively uncomplicated when we study *action* (such as acts of crime) because actions[1] are always taken by individuals (an aggregate cannot act[2]) and therefore action is linked to the individual (and to the individual as a unit of analysis). To study the influence of individual characteristics and experiences on action is therefore usually straightforward in regards to the basic unit of analysis (i.e., the individual).

However, the unit of analysis problem becomes much more complex when we introduce the *environment* and particularly when studying environmental influences on human action (such as acts of crime). There is no simple universal criterion to define the boundaries of 'the environment' (or of what it consists). The environment has to be defined in relation to something. What is a valid unit of environment depends on the nature of the research.

If the research question concerns the study of environmental influences on action, it is reasonable to argue that the valid unit of analysis is the part of the environment (the behavior-setting) in which an individual actually takes part. However, this is further complicated by the fact that individuals are not environmentally stationary but move around in space and are therefore likely to be *exposed* to a range of different kinds of settings (a generally neglected problem and one we will not deal with further in this chapter).

Most studies of the role of the environment in crime causation operate on an *aggregate* level, that is, the outcome variable represents aggregated data of individual acts of crime occurring in a geographic unit (and sometimes some or all of the predictors are aggregates of individual characteristics, such as mean income or percent from a particular ethnic background). Using aggregates of action as outcome variables (and, in addition, also using aggregates of individual characteristics as predictors) when trying to assess the environmental impact on individual action introduces numerous analytical problems.

To draw conclusions about causes of crime from aggregate associations between environmental characteristics and acts of crime, one has to be able to justify at least the assumption that the aggregate relationship holds up at the individual level (because, strictly speaking, it is not possible to argue for a causal relationship at the aggregate level, i.e., between aggregates). For example, if one studies the relationship between the level of informal social control and frequency of crime at the area level of analysis and finds a relationship between people's stated willingness to intervene to prevent crime (e.g., percent willing to intervene) and crime (e.g., mean

[1] Action consists of all movements of the body (e.g., talking or hitting) under the guidance of the individual (the latter implies the exclusion of reflexes).
[2] Although individuals can, of course, act jointly.

number of crimes committed per resident) it is not justifiable to argue that the *mean* is causally dependent on the *percent*. However, this association *may* reflect a causal relationship between individuals' exposure to settings with a certain level of informal social control and their decisions in those settings (as a consequence of the perceived level of informal social control) whether or not to commit acts of crime.

To avoid any misunderstandings, we do not argue that environmental (non-individual) conditions (such as opportunity or social climate) cannot be the causes of individual action. However, we do argue that ultimately it has to be shown (on the individual level) that such environmental conditions are directly linked to individual actions[3] (e.g., that individuals who commit more acts of crime *actually* have been exposed more often to settings with poor informal social control, and crucially, that their acts of crime *actually* have occurred in such settings).

A common problem in the study of aggregate relationships is knowing when it is justified to argue that an aggregate relationship reflects a (causal) relationship at the individual level. There are many important methodological issues to consider when inferring individual-level relationships from the study of aggregate data (a common problem in studies of environmental effects on acts of crime); although it is out of the scope of this chapter to deal with them all, we will mention a few.

In this chapter we will focus on *one* important problem: the size of the environmental unit of analysis. We advocate the (theoretical and empirical) advantages of using small-scale units of measurement when studying environmental factors in crime causation. We submit that this is important not only when analyzing data on the aggregate level (using aggregated data as outcomes) but also when analyzing corresponding relationship on the individual level. However, the empirical examples we use in this study only refer to analyses on the aggregate level and are selected to illustrate methodological points rather than present new substantive findings.

The dominant research tradition in the study of environmental influences on crime is the study of neighborhoods and crime. The methodological discussions in this tradition illustrate many problems related to the choice of units of analysis and, in our opinion, support our view that smaller is generally better.

## Neighborhood Studies of Crime and the Unit of Analysis Problem

Notwithstanding different theoretical approaches, various methodologies and often conflicting results, the terms 'community' and 'neighborhood' in criminological research generally imply that areas within cities (or somewhere on the urban-rural

[3] Although this is a minimal criterion for causal dependency, it does not imply that such a relationship is potentially causal. One must also provide a plausible causal mechanism that would explain *why* the environmental factor in question would cause (independently or in interaction) an individual to act in a certain way (e.g., to commit an act of crime). And of course, being able to manipulate putative causes and show that their outcomes vary as expected is the best way to demonstrate causal dependency (but rarely a realistic possibility when studying environmental influences on acts of crime).

continuum) constitute discernable spatial entities characterized by certain features of social organization which are relevant for behavior. While the collective social organization or ‘social climate’ of communities has always been theoretically considered a relevant characteristic of the behavioral context – for residents or visitors (Shaw and McKay 1942) – survey-based measurement of community social organization started much later, really progressing only during the last decade, both theoretically and methodologically (e.g., Kubrin and Weitzer 2003; Sampson et al. 2002; Wikström and Dolmen 2001).

In community or neighborhood-based studies, the problem of defining area size and boundaries is awkwardly inescapable and is typically solved pragmatically. Empirical studies usually depend on ‘official’ data regarding the socio-demographic and physical make-up of neighborhoods, and these data are normally available only for pre-defined administrative units; therefore, most studies simply use these administrative units. In the context of U.S. cities, these units represent street blocks, census blocks groups, or census tracts; in Britain, they represent electoral wards, enumeration districts, or, more recently, output (and super output) areas. Similar definitions and labels apply for other countries.

The problems connected with the choice of area units are well known in geography and other spatially oriented social sciences under the heading ‘modifiable area unit problem’ (MAUP) and have been debated extensively for decades. While a review of this debate is beyond the scope of this chapter (see further, Bailey and Gatrell 1995; Openshaw 1984; Openshaw and Taylor 1981; Reynolds 1998), it is useful to consider the two basic problems identified and their relevance to the study of the influence of environmental conditions on human action (such as acts of crime).

First, the ‘zonation effect’ relates to the difficulty of drawing meaningful boundaries within an area which reflect rather than blur the spatial patterns of important variables. For example, if an ethnic enclave is artificially cut in half by an administrative boundary such that the residents appear to live in two separate neighborhoods amongst the native population, indices of segregation will grossly underestimate the degree of segregation in this ethnic group. Social scientists using census and other official data are hardly ever in a position to change administrative boundaries. However, it seems fair to say that boundaries have often been intentionally defined to avoid such problems. The question then remains of how successful these attempts to preserve ‘natural’ patterns have been. Moreover, even if the boundaries follow ‘natural’ patterns for one dimension of segregation it is not certain it does so for other dimensions of segregation. However, when the research task is to study the influence by environmental conditions on human action, this problem is only relevant insofar as the boundaries of the units under study are drawn in such way so that they create large within-area heterogeneity in terms of the environmental conditions under study. Regardless of how boundaries are drawn, the smaller the units of analysis, the less likely it is that they will be significantly heterogeneous in their environmental conditions.

The ‘scale effect’ or ‘aggregation effect’ (or ‘aggregation bias’) concerns the susceptibility of statistical results to changes in the size of units. If the magnitude or even the direction of correlations between relevant variables depends on the level of spatial aggregation employed, results are less than robust and the question of which spatial level is the most appropriate for analysis then arises. Smith et al. (2000, p. 494), for example, discuss the possible interaction effects between routine activity variables and individual risk factors for victimization. A concentration of non-residential land use in one part of an area may be irrelevant for a household at the other end of this area if the area is large, or if activity patterns are constrained. McCord's et al. (2007) recent study shows that respondents' perceptions of neighborhood crime and disorder are in fact systematically linked to the distance from their household to non-residential areas within their neighborhood. On the whole, therefore, it seems fair to assume that smaller geographical units are more homogenous, and hence more accurately measure environments. In other words, smaller is better.

Contrary to the 'zonation effect', researchers can often choose between and compare results using different spatial levels such as census blocks versus census tracts. Some studies have investigated the 'scaling effect' or 'aggregation bias' by comparing statistical analyses of socio-demographic factors and crime rates on two levels of aggregation. Ouimet (2000) compared census tracts (averaging 3,500 inhabitants) to groups of census tracts ('neighborhoods' averaging 21,000 inhabitants) in an aggregate-level analysis of crime rates in the city of Montreal (Canada). He found higher bivariate correlation coefficients, beta-coefficients, and R-squares for the larger neighborhoods which he suggests is due to inflated random variation in the smaller census tracts due to low absolute numbers of crimes. However, some associations, such as that between subway stations and violent juvenile crime, are stronger and more significant for the smaller census tract level. The effect of land use patterns on crime may be attenuated by aggregating small area units to larger but more heterogeneous units as discussed by Smith et al. (2000). Ouimet's (2000) decision to use larger 'neighborhood' units seems misguided, because by looking exclusively at *standardized* coefficients, which reflect the amount of variation around a regression line, he misses the likewise important information contained in *un*standardized coefficients, i.e., how strong a predicted effect is in terms of the change of units. We will demonstrate this point in the empirical part of this chapter.

Wooldredge (2002) reported a similar comparison based on data from Cincinnati (USA), which is divided into 129 census tracts and 48 neighborhoods. He employed multilevel analysis, entering individual data on arrestees on the first level and contextual data on the second (tract or neighborhood) level. Two significant area-level effects were rendered insignificant when switching from the smaller tract to the larger neighborhood level. However, further tests revealed that the coefficients were not significantly different. Wooldredge concluded that there were no substantial differences between aggregation levels apart for the effects of aggregate sample sizes (Wooldredge 2002, p. 699).

In the light of these results, Sampson's (2006, p. 35) comment that 'empirical results have not varied much with the operational unit of analysis' and that the social stratification of communities is 'a robust phenomenon that emerges at multiple levels of geography' seems warranted. In fact, many survey-based studies employ relatively large units of analysis such as census tracts (US) or wards (UK) which typically encompass 4,000–10,000 residents (e.g., Bellair 1997; McVie and Norris 2006; Sampson et al. 1997). Only some studies, such as the Seattle Victimization Survey (Miethe and Meier 1994) use smaller units, mainly street blocks. In Ralph Taylor's (2001) Baltimore study, some analyses are carried out on the street block level, although the main focus remains on census tracts.

## Statistical Power Considerations in Multi-level Sampling Designs

There is an additional reason why the 'unit of analysis' issue is of particular relevance for survey-based community studies on crime causation. In contrast to studies employing recorded crime data, which usually represent complete samples, survey-based community studies using random samples of respondents are subject to the problems of statistically inferring relationships using standard errors and significance levels. Considerations of statistical power for hypothesis testing are more complicated in multi-level studies where respondents are clustered in neighborhoods (or other social groups) and hypotheses refer not only to individual-level effects between respondents, but also to aggregate-level effects between neighborhoods or cross-level effects between individuals and neighborhoods.

Recent methodological research has advanced knowledge concerning statistical power analysis in complex survey designs (Murrey et al. 2004; Raudenbush 1999). A detailed technical discussion of this research is beyond the scope of this chapter; however, it is important for any community-based study working on a restricted budget to balance the number of neighborhoods, and the number of individual respondents within neighborhoods, surveyed in order to achieve an optimal statistical power to test hypotheses of interest (Snijders and Bosker 1999, p. 140). Consider a study on community-level effects on crime which aims to measure the social organization of all neighborhoods in a given city. Assume, for example, that this study has resources to survey 1,000 respondents. These respondents may theoretically be allocated to 10 neighborhoods with 100 respondents each, or to 100 neighborhoods with 10 respondents each, or to any other combination of areas and individuals within areas, depending on the existing levels of administrative units in the city. Statistical power analysis for this task is complex and depends on the focus of the hypotheses. However, generally and within certain limits, more statistical power is gained by choosing more areas with fewer respondents rather than fewer areas with more respondents (Murrey et al. 2004, p. 424; Snijders and Bosker 1999, p. 152). This is mainly due to the large incremental increase in statistical power for detecting significant area-level effects if the initial number of areas is small. As Snijders and Bosker (1999, p. 140) remark, 'requirements on the sample size at the highest level... are at least as stringent as requirements on the sample in a single level design'. Ten, fifty, or even hundred aggregate-level units still represent a small sample size for multivariate analyses. Using simulation studies, Snijders and Bosker (1999, p. 152) show that depending on the focus of the hypotheses, as few as 8–15 respondents per area are sufficient to achieve statistical significance in multi-level models including area-level effects.

Taking together the issue of homogeneity of small areas and power considerations for area-level sample size, we are not satisfied with Wooldredge's and Sampson's conclusions that area size does not matter. Instead, we posit that using smaller units of analysis has important advantages over using larger units. First, the studies just mentioned clearly evidence differences, even if these were marginal. Second, both studies started with units of analysis which were already quite large, and therefore insufficiently homogenous for the measurement of behavioral contexts. Hence, the crucial question is whether geographical areas large enough to encompass 5,000–8,000 inhabitants can approximate behavior-settings which influence human development and action in any meaningful sense. We would argue they cannot. Thus, we deem it worth digging deeper into the issue of scale effects in a multilevel framework which employs very small units of analysis averaging only 300 residents.

## The Peterborough Adolescent and Young Adult Development Study (PADS+)

*The Peterborough Adolescent and Young Adult Development Study* (PADS+) is an ongoing Economic and Social Research Council (ESRC) financed longitudinal study of young people's development and crime involvement during adolescence and young adulthood. The study includes 716 subjects randomly selected from a cohort of young people who were 12 years old in 2003 and living in the city of Peterborough (UK) or several surrounding villages.

The overall aim of PADS+ is to contribute to a better understanding of the causes and prevention of young people's crime involvement by studying (i) the interaction between individual characteristics and experiences and the features of the social environments in which young people develop and act and (ii) how these interactions change and shape criminal involvement over two critical developmental phases: adolescence and the transition into young adulthood.

The data collected in the Peterborough longitudinal study (PADS+) covers three main topics: (i) *the individual*: his and her individual and social characteristics and experiences (data is collected through an interview, interviewer-led questionnaires and psychometric tests); (ii) *the environment*: the characteristics of different small-area environments of Peterborough (data is collected through a community survey); and (iii) *individuals' exposure to different environments* in Peterborough (data is collected using a Space-Time Budget technique).

The study has to date (2008) successfully completed five waves of data collection, with a 98% retention rate by wave five. Data was also collected via a special community survey in 2005 which covered a random sample of the Peterborough population and received responses from approximately 6,600 residents. The purpose was to gather detailed information at a small area level (approximating behavior-settings) about key environmental characteristics within the city and nearby villages of Peterborough. The data used in this chapter is taken from the 2005 Peterborough Community Survey (PCS)

## The Community Survey Data

The PCS was conducted in 2005. The sample is a clustered random sample of adult residents living in the city of Peterborough and several adjacent villages. Peterborough has approximately 160,000 inhabitants. Respondents aged 18 and over were randomly selected from the electoral register, which comprises 104,281 eligible voters; 4.9% of the adult population in Peterborough is not included in the electoral register, mainly because of non-British citizenship. In addition, 35% of registered voters 'opt out' which means their address cannot be used for purposes like surveys. It is generally assumed that those who opt out are more likely to be middle class and have a better than average education.

The PCS used the smallest available administrative units called 'output areas' (OAs). OAs have been empirically derived using individual-level census data and geographical and physical information to approach homogeneity both in terms of socio-demographic composition and population size (Martin 2000). OAs, on average, have around 300 inhabitants. This means that in densely populated areas there are more smaller OAs than in sparsely populated areas.

We grouped all OAs according to the official 'Index of Multiple Deprivation' (IMD) into 'normal' and 'deprived' OAs. The IMD is an overall index measuring deprivation in seven domains, including income, employment, education, and health

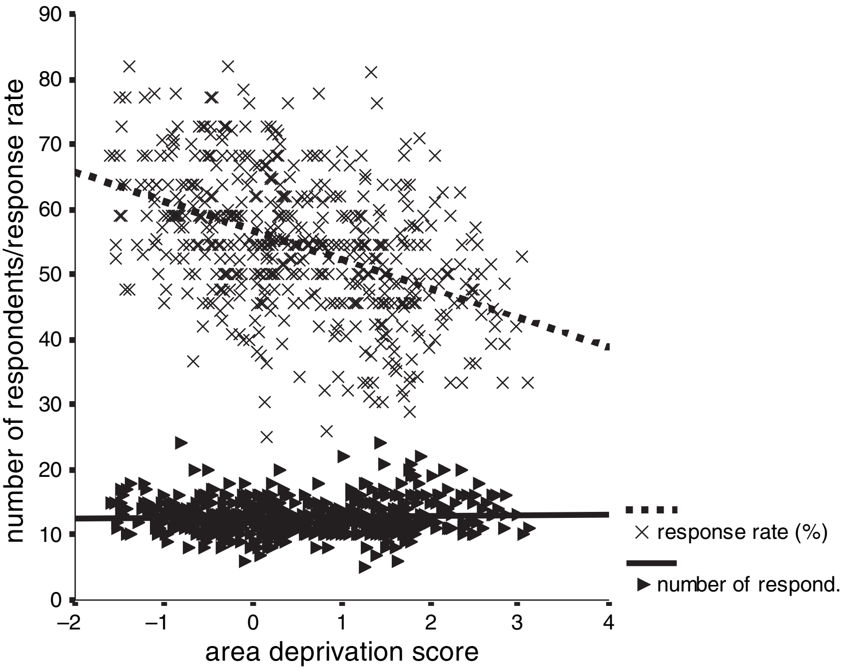


**Fig. 2.1** Absolute number of respondents and response rates by area deprivation (N = 518 output areas with N = 6, 615 respondents)

(Office of the Deputy Prime Minister 2004). Anticipating a lower response rate in deprived areas, we randomly selected 22 addresses from ‘normal’ OAs and 33 addresses from ‘deprived’ OAs. Respondents were sent a 20-page questionnaire and, if needed, a reminder letter. In OAs where we initially received less than 15 respondents, we sent up to four additional reminder letters. The overall response rate was 53%. By oversampling respondents in deprived OAs and sending up to four reminder letter to non-responding persons in OAs which lacked a sufficient number of respondents, we counter-balanced the well-known effect of low response rates in deprived areas. As can be seen in Fig. 2.1, this strategy worked well in the sense that the absolute number of respondents does not co-vary with area deprivation. The average number of respondents per OA was 12.8 (standard deviation 2.7).

In the questionnaire, we introduced questions on the respondents’ residential areas as follows: ‘We would like you to think about the area within a short walking distance (say a couple of minutes) from your home. That is the street you live in and the streets, houses, shops, parks and others areas close to your home’. We intended for this definition to focus respondents’ answers on the immediate area around their homes (approximating the size of a behaviour setting).

## Analytic Strategy

By asking respondents questions about their immediate area of residence, we intend to measure collective properties of these local contexts. As with any social science measurement, however, respondents’ answers will not represent a ‘true’ picture of their area but will be biased and affected by measurement error to some extent. For example, a respondent who spends little time in his or her immediate neighborhood and does not care about his or her neighbors will probably know less about community life than respondents who spend more time in their immediate neighborhood and have more contacts with their neighbors. A very old respondent may evaluate victimization risks differently from a younger respondent or may report less alcohol-related disorder because he or she seldom leaves his or her home at night. Some aspects of community life, such as social cohesion and trust among neighbors, are more subjective and depend on individual experiences, while others, such as the presence of litter in front gardens, are more objectively measurable, and depend less on (but are not entirely independent of) respondents’ individual characteristics.

We start with some simple examples to illustrate the accuracy of responses by exploiting external spatial information. In the questionnaire, respondents were asked to state whether certain aspects of infrastructure like shops, police stations, etc. were close to their homes. Figure 2.2 shows the percentage of respondents who reported the presence of a police station in their local area as well as the actual location of police stations (symbolized by stars). The close congruence, which was also found for fire stations (graph not shown), shows that people were quite accurate in reporting spatial aspects of infrastructure which we are able to check. We also found, for example, a close spatial match between survey respondents’ frequency

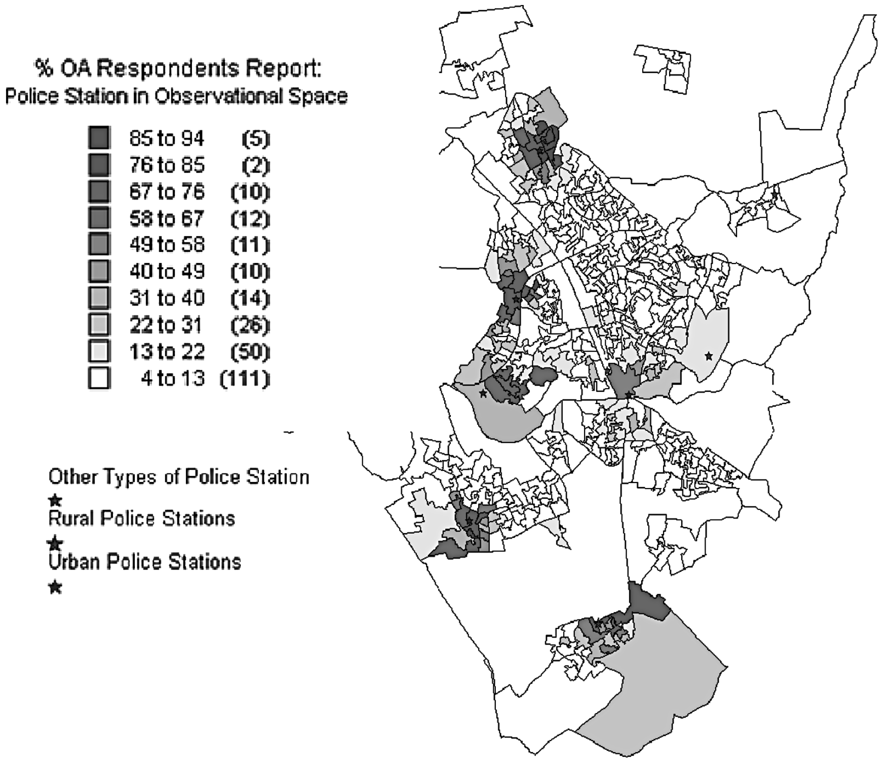


**Fig. 2.2** External validation of respondents' assessment of the vicinity of police stations

of reporting noisy neighbors and the frequency of police calls regarding nuisance neighbors (finding not shown). These and other similar tests make us confident that the respondents reporting of their observable local conditions are largely accurate. However, in matters concerning latent dimensions of social organization, the same opportunity for external validation does not exist.

The question then remains as to how we can evaluate the quality of survey data on contexts and determine the degree to which respondents' answers reflect the actual social conditions of their common environment rather than their subjective views. As in any kind of empirical research, data quality refers to validity and reliability. Validity is the degree to which data actually measures what it is supposed to measure, while reliability is the degree to which a measure is consistent (or precise).

Until recently, the reliability of community-level survey data has been generally neglected, as no standardized statistical tools comparable to scale reliability measures like Cronbach's alpha existed. This has changed largely thanks to the 'ecometric' approach recently developed by Raudenbush and Sampson (1999) which is specifically designed to assess the reliability of data on collective entities, such as neighborhoods, schools, or companies. This statistical method is based on the idea, analogous to the concept of interrater reliability, that information given by individuals concerning a common environment is reliable to the extent that it is concurrent. If all respondents in a neighborhood answer questions identically, their information is deemed perfectly reliable. On the other hand, if every respondent gives a different answer, they do not provide a consistent picture of neighborhood conditions. Raudenbush and Sampson call the consistency of residents' reports on their shared environment '*ecological reliability*'.

The same approach can be exploited to address the issue of homogeneity within area units. If smaller output areas are more homogeneous in terms of environmental conditions than larger super output areas (which represent clusters of smaller output areas), then we can expect answers (the subjects' observations) on the smaller spatial level to be more similar, and hence more reliable.

Statistically, the measurement of ecological reliability is based on multilevel (or hierarchical linear) modeling (Hox 2002; Raudenbush and Bryk 2002). One of the basic features of multilevel modeling is a decomposition into within- and between-group variance, where the share of the between-group variance represents the degree of consistency of answers by members of the same group. The more concurrent the answers from respondents in one area are, the lower the within-group variance, and the higher the between-group variance. Computed from this variance decomposition, the intra-class correlation coefficient (ICC) is defined as the share of the between-group variance of the sum of between- and within-group variance. The coefficient of ecological reliability called lambda is based on the intra-class coefficient weighted by the number of respondents, just as Cronbach's alpha is weighted by the number of items in a scale. Thus, the ecological reliability increases with the number of respondents. Finally, using a Bayesian approach, estimates of group-level values are 'smoothed' by pulling them towards the mean of all groups if the reliability is low due to very few respondents (see below).

One of the important findings from Raudenbush and Sampson's research is that a relatively small number of respondents are sufficient to achieve a high reliability of information on neighborhood conditions. Little incremental improvement of the ecological reliability is observed beyond 30 or 40 respondents (Raudenbush and Sampson 1999, p. 9). The fact that ecological reliability is dependent on the number of respondents is particularly critical in the case of the PCS considering its relatively low number of respondents per small OA. It is important to determine if an average of 13 respondents is enough to achieve a reliable measurement of neighborhood characteristics.

On the following pages, we will use the ecometric approach to assess the reliability of survey scales measuring behavioral contexts in Peterborough comparing smaller and larger area units. We will try to answer to what extent respondents' observations represent concurrent and reliable views on their local area and how much internal homogeneity is dependent on the level of aggregation. The ecometric approach will enable us to assess the advantage of our sampling design based on many very small units compared to more conventional designs which use fewer but larger units. Subsequently, we will then investigate how both levels of aggregation compare in multi-level regression models which include area-level effects. Here, the important question is whether the increase in area-level sample size enhances the statistical power to detect significant effects.

We start with the smaller spatial units of output areas; in a second step, we compare these results with those from larger spatial units (super output areas) and try to assess the advantages and disadvantages of our focus on the smaller output area level. We use the 'social cohesion/trust' and the 'informal social control' scales which jointly constitute the 'collective efficacy' scale as an example. Incidentally, the collective efficacy scale can theoretically be viewed as a measure of an area's moral context.

## Results: 'Social Cohesion/Trust' and 'Informal Social Control' (Collective Efficacy)

Collective efficacy has emerged as an important dimension in recent community research, consisting of questions about trust and cohesion among neighbors (social cohesion) and the capability of neighbors to exert informal social control over misbehaving children and adolescents (informal social control). Sampson et al. (1997) argued, in their seminal paper on the attenuating effects of collective efficacy on violence in Chicago neighborhoods, that both aspects of collective efficacy – trust among neighbors and their shared expectation about counteracting disorder and crime – are so closely associated theoretically and empirically to justify a unified concept. Other studies and analyses keep these subscales separate or focus on informal social control (Wikström and Dolmen 2001; Silver and Miller 2004; Taylor 2002). We will do likewise, assessing each scale separately and look to informal social control as the dependent variable in the final analysis.

The first four items displayed in Table 2.1 were introduced in the questionnaire as follows: 'For each of the following, please state if it is very likely, likely, unlikely, or very unlikely that people in your neighborhood would act in the following manner.' Respondents were asked about the likelihood 'that your neighbors would do something about it' if neighborhood children were skipping school, hanging out on a street corner or spraying graffiti on a local building or, if they were fighting.

Confirming the results of previous studies, principal component analysis with oblimin rotation of all nine items yields two closely related dimensions representing social cohesion/trust and informal social control (Table 2.1). The two dimensions are clearly correlated ($r = 0.45$ on the individual level, $r = 0.77$ on the area level). The individual-level reliabilities are very high for both sub-scales of collective efficacy (Cronbach's alpha $= 0.83$).

To evaluate the ecological, neighborhood-level reliability of these scales, the first step within multilevel modeling is to compute ICCs in a so-called 'empty model' without any individual-level predictors, comparable to variance decomposition in a conventional analysis of variance. As reported in Table 2.2, about 19% of the variance of 'social cohesion/trust' and about 11% of the variance of 'informal social control' is due to differences between output areas. Weighted by the number of

**Table 2.1** Principal component analysis of 'collective efficacy' scale (individual level, N = 6, 615 respondents)

| Variables | Dimensions: 'informal social control' (Eigenvalue 4.4, 48.5% variance) | 'social cohesion' (Eigenvalue 1.3, 14.5% variance) | Extraction |
|---|---|---|---|
| Skipping school and hanging out | 0.82 | | 0.67 |
| Spray-painting graffiti on a local building | 0.82 | | 0.71 |
| Fight in front of your house | 0.77 | | 0.59 |
| Child showing disrespect to an adult | 0.81 | | 0.64 |
| People around here are willing to help their neighbours | 0.34 | 0.57 | 0.62 |
| This is a close-knit neighbourhood | 0.30 | 0.58 | 0.58 |
| People in this neighbourhood can be trusted | | 0.74 | 0.67 |
| People in this neighbourhood generally do not get along with each other | | –0.84 | 0.63 |
| People in this neighbourhood do not share the same values | | –0.81 | 0.58 |
| Cronbach's alpha | 0.83 | 0.83 | |

**Table 2.2** Variance components of 'social cohesion/trust' and 'informal social control' scales (N = 518 OAs with N = 6, 615 respondents)

| | Social cohesion/trust | | Informal social control | |
|---|---|---|---|---|
| | Empty model | Conditional model[a] | Empty model | Conditional model[a] |
| *variance components* between respondents ($r_{ij}$) | 0.41126 | 0.40716 | 0.81604 | 0.81235 |
| explained variance on respondent level | – | 1.0% of 0.41126 | – | 0.5% of 0.81604 |
| between areas ($u_{0j}$) | 0.09878 | 0.0891 | 0.10425 | 0.10042 |
| explained variance on output area level | – | 9.8% of 0.09878 | | 3.7% of 0.10425 |
| ICC[b] | 19.4% | 18.0% | 11.3% | 11.0% |
| lambda ('ecological' reliability) | 0.75 | 0.72 | 0.61 | 0.60 |

[a] *Controlling for socio-demographic composition*
[b] *Intraclass correlation coefficient* $= r_{ij}/(u_0 + r_{ij})^{*}100$

respondents, this ICC translates to lambdas of 0.75 for 'social cohesion/trust' and 0.61 for 'informal social control'. Whereas the result for the former scale is very good, the value for the latter is at best satisfactory. We can interpret these reliability measures as showing that respondents are more concurrent in giving their impression of social cohesion and trust in their area of residence than assessing the likelihood that their neighbors would intervene in situations of child and adolescent misbehavior.

In the Chicago Project on Human Development in Chicago Neighborhoods' (PHDCN) community survey (which uses substantially larger area units than this study), the ICCs reported for the same scales were slightly higher, at 24% for social cohesion/trust and 13% for informal social control (Raudenbush and Sampson 1999, p. 8). In a community survey in two German cities using the same scales (but, again, larger area units), the ICC was 17% for 'social cohesion/trust' and 10% for 'informal social control', mirroring findings from Peterborough (Oberwittler 2003).

In a second step, individual-level variables are introduced to the multilevel model to control for neighborhoods' socio-demographic composition. If differences in measurement between neighborhoods are mainly due to socio-demographic variables, for example, the fact that older respondents answer survey questions differently than younger respondents, or poorly educated respondents answer differently than highly educated respondents, this would be reflected in the so-called conditional model. The stronger the effect of individual-level variables, the more the ICC would be reduced between the empty and conditional model. The second column in Table 2.2 ('conditional model') reveals that socio-demographic variables have only a very small effect on area-level measurements, reducing the ICC of 'social cohesion/trust' by 10% and the ICC of 'informal social control' by 4%. Socio-demographic variables also have a very marginal effect on within-area differences between individual respondents (1% for 'social cohesion/trust', 0.5% for 'informal social control'). This result underlines the fact that perceptions of area social organization are largely independent of individual socio-demographic factors, and thus area differences are hardly attenuated if controlling for respondents' socio-demographic composition. What really drives area differences in these dimensions of collective efficacy, then, is the effect of the *collective* makeup of areas, such as the concentration of social deprivation. We will turn to these collective dimensions later.

In the case of Peterborough, an average of four to five output areas constitutes one super output area. Any additional information gained on the smaller level can be assessed by aggregating the survey data to the super output area level and comparing the results from both (Fig. 2.3). If we look to the ecological reliability of the social cohesion/trust scale, its ICC (share of between-group variance) is only 14% on the superoutput areas (SOA) level compared to 19% on the smaller OA level, indicating that SOAs are internally more heterogeneous.[4]

On the other hand, because there are more respondents per SOA unit, the ecological reliability lambda rises from 0.75 to 0.91. The same holds true for the 'informal social control' scale. The ICC of 'informal social control' drops from 11 to 8.5% if

[4] However, it should be noted that the super output areas are still generally smaller (averaging about 1,500 inhabitants) than the units commonly used in ecological studies. We would therefore expect that the difference in heterogeneity would be even greater had our comparison with OAs involved larger units than the SOAs.

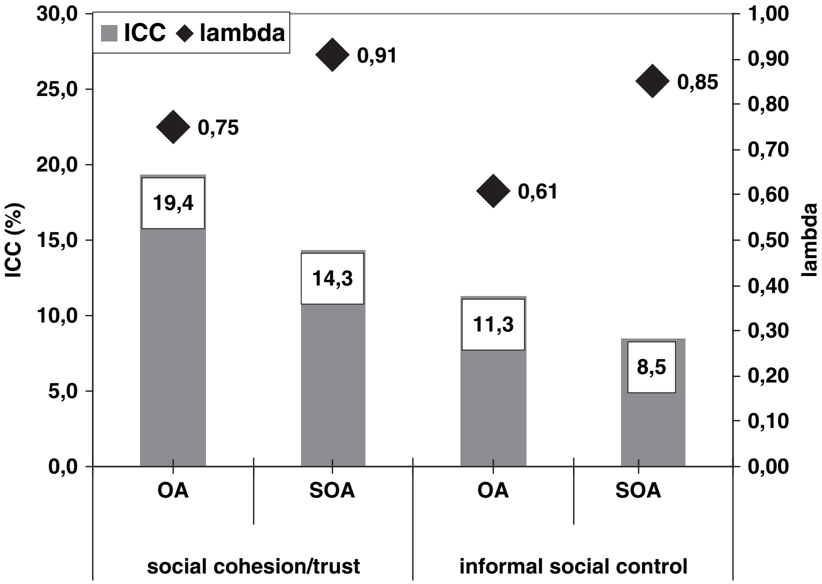


**Fig. 2.3** ICCs and ecological reliability (lambda) of 'social cohesion/trust' and 'informal social control', OA and SOA levels compared (N=518 OAs, N=90 SOAs with N=6, 615 respondents)

one move from OA to SOA level; lambda, on the other hand, increases from 0.61 to 0.85.

Thus, one faces a trade-off between the (theoretically important) aim to target small and homogenous areas approximating behavior-settings and the need for sufficiently reliable measurements (methodologically important). The mean of 13 respondents per output area certainly marks the lower bound of a reasonable sampling design. In the Chicago survey (PHDCN study), using much larger area units, the average number of respondents per neighborhood cluster was 26 (Raudenbush and Sampson 1999). As we will see below, a rather unintended positive side-effect of lower reliability and larger measurement errors is the reduction in multicollinearity between neighborhood-level variables.

We can achieve a more detailed picture on the gains in homogeneity on the smaller OA level by utilizing multilevel analyses to build three-level models where the variance is more accurately decomposed into shares of variance for each level. In this model, level 1 represents respondents nested in OAs, level 2 represents OAs nested in SOAs and level 3 represents SOAs. Figure 2.4 displays the results of this variance decomposition. For both scales, the larger share of between-group variance is between the SOAs, with around 30% of the additional variance lying between the OAs within these SOAs. Thus, the heterogeneity between OAs within one SOA is lower than the heterogeneity between SOAs. Still, there is a considerable increase of around 30% in spatial homogeneity.

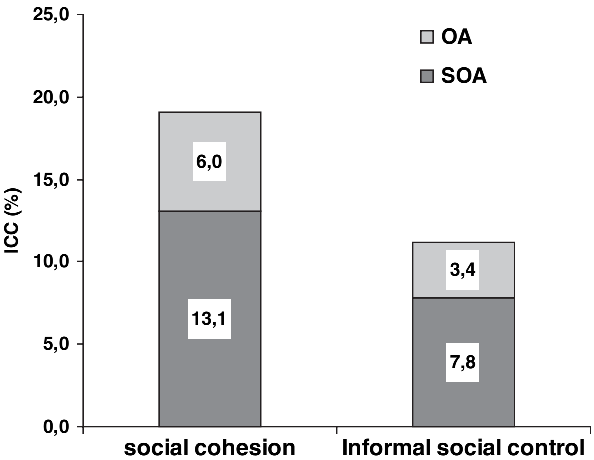


**Fig. 2.4** ICC's of 'social cohesion/trust' and 'informal social control' in three-level models N = 6,615 respondents nested in N = 518 OA's nested in N = 108 SOA's)

## Multiple Regression Models with Area Level Predictors

How does the decision to use 518 smaller instead of 102 larger spatial units change the results of substantive statistical analyses? We will explore this in this final section by modeling stepwise multiple regression models on both levels of aggregation with informal social control as the dependent variable (Tables 2.3 and 2.4). Models on both levels are kept as similar as possible in order to facilitate the comparison and to focus on the 'scaling effect'. These models are pure area-level models because we are dealing with area-level dimensions of social structure (measured by census data) and social organization (measured by survey data). However, all survey scales are empirical Bayes estimates (which 'smoothen' unreliable values at the extreme ends of the distribution) adjusted for individual-level socio-demographic composition, as recommended by Sampson et al. (1997).

The models are built in three steps. First, only structural (census) variables are used to predict the outcome. In a second step, observed disorder and area neighbor contacts are introduced, which are assumed to correlate negatively (disorder) respectively positively (contacts) with expectations for informal social control. In the final step, the predictive power of the second component of collective efficacy, social cohesion/trust, is also tested. The 'R squared change' value indicates how much the model improves with each step. We report unstandardized coefficients which can be easily compared as all predictors have been standardized to a mean value of 0 and a standard deviation of 1.

Comparing the models at the two levels of aggregation, the first thing which is clear is that, on the whole, the results are concurrent. No predictor behaves in a completely different manner or changes signs when moving from the smaller OA to

**Table 2.3** Aggregate-level OLS regression of 'informal social control' (N = 518 OAs with N = 6, 615 respondents)

| | **1** | | **2** | | **3** | |
|---|---|---|---|---|---|---|
| *N = 518 OAs* | unst. B | t-value | unst. B | t-value | unst. B | t-value |
| **Constant** | 3.172 | 323.8 | 3.132 | 333.495 | 3.106 | 362.911 |
| deprivation (factor score) | **−0.096** | **−9.896** | **−0.040** | **−4.008** | −0.011 | −1.192 |
| asian ethnicity (factor score) | **−0.034** | **−2.753** | −0.012 | −1.126 | 0.014 | 1.405 |
| asian ethn. square | **0.006** | **3.639** | 0.002 | 1.728 | −0.001 | −.775 |
| fluctuation | **−0.020** | **−2.041** | −0.014 | −1.675 | 0.003 | 0.399 |
| fluctuation * deprivation | 0.015 | 1.681 | 0.011 | 1.374 | 0.007 | 0.942 |
| fluctuation ∗ asian ethn. | **−0.015** | **−2.147** | −0.010 | −1.604 | −0.006 | −1.192 |
| pop density | **−0.070** | **−5.699** | **−0.037** | **−3.349** | −0.016 | −1.597 |
| non-residential land use | **−0.035** | **−3.430** | **−0.018** | **−2.034** | −0.015 | −1.909 |
| ***R square*** | ***0.45*** | | | | | |
| observed (youth) disorder | – | | **−0.118** | **−11.110** | **−0.072** | **−7.131** |
| neighbourh contacts | – | | **0.041** | **5.135** | 0.011 | 1.518 |
| neighbourh contacts ∗ depriv. | – | | −0.012 | −1.651 | **−0.016** | **−2.359** |
| neighbourh contacts ∗ asian ethn | – | | **0.015** | **2.776** | **0.017** | **3.570** |
| ***R square change*** | | | ***0.13*** | | | |
| social cohesion | – | | – | | **0.132** | **12.085** |
| ***R square change*** | | | | | ***0.09*** | |
| ***Total adj. R square (F-value)*** | ***0.44*** | ***(51.5)*** | ***0.58*** | ***(58.2)*** | ***0.68*** | ***(80.4)*** |

**bold** coefficients: $p < 0.05$; all predictors are z-standardized

**Table 2.4** Aggregate-level OLS regression of 'informal social control' (N = 102 SOAs with N = 6, 615 respondents)

| | **1** | | **2** | | **3** | |
|---|---|---|---|---|---|---|
| *N = 102 SOAs* | unst. B | t-value | unst. B | t-value | unst. B | t-value |
| **Constant** | 3.231 | 170.520 | 3.155 | 153.702 | 3.109 | 155.439 |
| deprivation (factor score) | **−0.175** | **−9.055** | **−0.074** | **−3.051** | −0.026 | −1.096 |
| asian ethnicity (factor score) | **−0.072** | **−2.638** | −0.019 | −0.767 | 0.033 | 1.344 |
| asian ethn. square | **0.010** | **2.957** | 0.004 | 1.102 | −0.002 | −0.761 |
| fluctuation | −0.017 | −1.083 | −0.019 | −1.361 | −0.003 | −0.252 |
| fluctuation * deprivation | −0.015 | −0.805 | −0.021 | −1.331 | −0.011 | −0.811 |
| fluctuation * asian ethn. | −0.012 | −1.018 | 0.000 | 0.038 | −0.006 | −0.657 |
| pop density | **−0.042** | **−2.260** | **−0.038** | **−2.340** | −0.015 | −0.966 |
| non-residential land use | −0.014 | −0.768 | −0.018 | −1.111 | −0.021 | −1.542 |
| ***R square*** | ***0.76*** | | | | | |
| observed (youth) disorder | | | **−0.129** | **−5.614** | **−0.083** | **−3.756** |
| neighbourh contacts | | | 0.024 | 1.785 | 0.005 | 0.416 |
| neighbourh contacts * depriv. | | | −0.004 | −0.260 | 0.002 | 0.133 |
| neighbourh contacts ∗ asian ethn | | | 0.020 | 1.856 | 0.012 | 1.260 |
| ***R square change*** | | | ***0.07*** | | | |
| social cohesion | | | | | **0.146** | **5.261** |
| ***R square change*** | | | | | ***0.04*** | |
| ***Total adj. R square (F-value)*** | ***0.74*** | | ***0.81*** | | ***0.86*** | |

**bold** coefficients: $p < 0.05$; all predictors are z-standardized the larger SOA level (see Table 2.4). Thus Sampson's assertion that most results are robust across level of aggregation is at least supported by these empirical findings. Yet differences are noticeable in the more subtle parts of the models, mainly concerning the moderate and weak predictions. For example, in the first step there are significant but weak effects of residential stability and land use at the OA level which are insignificant at the SOA level. There is also a significant interaction effect between residential instability and Asian ethnicity, suggesting that negative effects on informal social control are exacerbated if both dimensions go hand in hand. In the second model, the significant effect of neighborhood contacts as well as its interaction term between Asian ethnicity disappears when moving to the larger spatial level.

It appears, then, that there is a general tendency for weaker, more subtle effects to disappear when data in analyzed at a higher level of aggregation. This could be due to watering down the degree of spatial homogeneity by aggregating small areas. For example, if there are small ethnic enclaves or 'pockets' of non-residential land use, their effect may be masked to the extent of non-significance if they are lumped together with neighboring areas. More detailed research into geographical micro-spaces would be needed to elaborate this hypothesis.

There is another, more technical reason why regression models on the lower level of spatial aggregation yield more nuanced and complex findings. The statistical power to detect significant effects of predictors necessarily increases with the sample size of area units, which renders the standard errors of coefficients smaller. This is reflected by higher t-values in the OA models reported in Table 2.3 compared to the SOA models in Table 2.4. Logically, it is the more subtle effects on the lower end of significance – like interaction effects typically are – which profit from this increase in power. An important finding, therefore, is that a sufficient number of area units are an important requirement for more complex modeling.

Finally, the share of explained variance (R squared) is considerably higher in models on the larger SOA level than on the smaller OA level. This should, however, not be interpreted as a decisive advantage, as Ouimet (2000) did, since the unstandardized coefficients are not affected by spatial aggregation. This effect can be graphically illustrated in the scatterplots displayed in Fig. 5a and 5b. In both scatter plots, the association between deprivation and expectations for informal social control are overlaid for both spatial levels, where every dot represents one area. The steepness of the regression line (the slope) represents the strength of the effect of deprivation on informal social control. In Fig. 2.5a, the lines are virtually the same, yet the cloud of dots is much more widely dispersed around the regression line on the OA level; there are some particularly extreme outliers on both ends of the scale which 'disappear' when the data is aggregated to the larger SOA level. It is important to understand that although the correlation coefficient and the R squared value are necessarily much higher at the SOA level due to less random variation, the predicted effect of deprivation on informal social control (the steepness of the regression line, formally expressed by the *unstandardized* beta coefficient) is practically the same. This is the reason why a focus on correlation coefficients, standardized regression coefficients, and value of R squared often is unhelpful and even misleading, and why

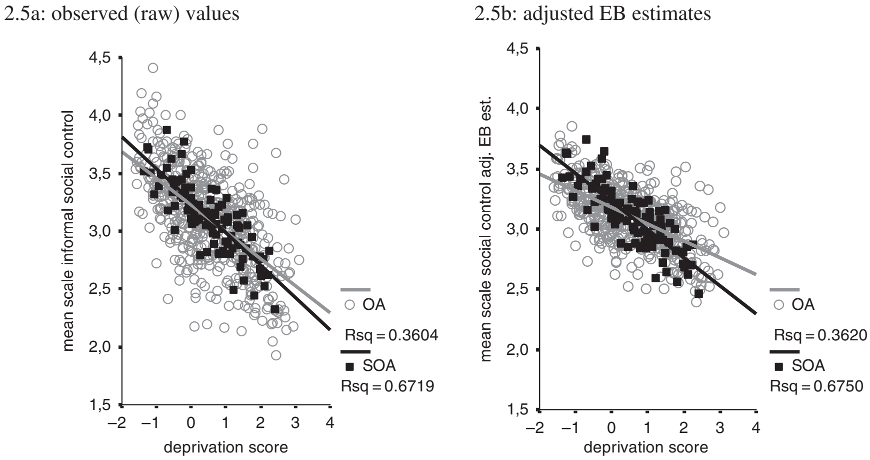


**Fig. 2.5 a,b** Scatter plots of the association between area deprivation and informal social control, OA and SOA levels compared (N = 518 for OAs, N = 108 for SOAs and N = 6,615 for respondents)

the information carried in the unstandardized beta coefficient is more meaningful if one is interested in prediction.

However, the Bayesian approach to the estimation of group means does impact the strength of the relationships at the smaller OA level. The scatter plot on the right (Fig. 2.5b) displays the same association using the empirical Bayes estimates adjusting for individual socio-demographic composition which we used in the regression models. As one can easily see, in comparison to the 'raw value' on the left side, the Bayes method 'pulls' the extreme values at both ends of the distribution towards the mean for all neighborhoods, rendering the distribution of the OA level group means more similar to that of the SOA level. This happens to the extent that outlying estimates of group means are deemed unreliable, which has been shown to be a function of low sample sizes of respondents within areas. As result of this, the regression line is rendered flatter than that of the scatter plots using either the raw or the SOA values. In effect, the empirical Bayes 'smoothing' procedure attenuates the empirical association and proposes a more conservative estimate. This can be observed in the regression models discussed above where the unstandardized coefficient of deprivation is $-0.175$ at the SOA level but only $-0.096$ at the OA level. Generally speaking, if the number of respondents per spatial unit is very small and estimates of survey results become too unreliable, the Bayesian method built into multi-level modeling software will tend to produce very conservative estimates which may in extreme cases effectively 'kill' substantial results. However, the comparison of models shows that, on balance, more significant effects are observed at the lower spatial level because the effect of increased statistical power due to a larger number of area units is stronger than the effect of attenuated reliability of area estimates due to the smaller number of respondents in areas.

There is yet another side-effect of using smaller spatial units which enables more nuanced and complex models and the testing of more differentiated hypotheses. At the lower spatial level of OAs, there is a much higher degree of randomness or random variation due to the 'small number problem' where a change of one or two individual values can cause a significant change in rates. While this random variation or random noise could be viewed as a nuisance, it also has the paradoxically positive effect of attenuating the problem of multicollinearity which hampers aggregate data analysis in particular (Land et al. 1990). If independent variables are strongly intercorrelated, with bivariate correlations greater than $r = 0.70$, the basic task of multiple regression analysis (to disentangle the effects of intercorrelated variables) becomes very difficult. From this perspective, using data on a smaller level of aggregation, where correlations are generally lower, can actually be advantageous (Sampson and Raudenbush 1999, p. 625). At the SOA level, correlations between deprivation, informal social control and social cohesion all surpass 0.80, whereas they range between 0.60 and 0.80 at the OA level.

## Conclusions

In this chapter, we have made a number of arguments about how to advance the study of the role of the social environment in crime causation. We have, on theoretical grounds (based on Situational Action Theory), criticized the common practices of (i) using large area units and (ii) neglecting the importance of individuals' exposure to different kinds of environments (within and outside their neighborhoods). We have advocated the use of small area units (which resemble behavior-settings as closely as possible) combined with a measure of individuals' exposure to different behavior-settings (within and outside their neighborhoods) to advance empirical study of the role of the social environment in crime causation. In this chapter, we focused on exploring the pros and cons of using small area units. The equally important problem of measuring individuals' exposure to different environments will be dealt with elsewhere.

First, we have shown that differences in subjects' assessments of the environment of their immediate area of residence (e.g., social cohesion, informal social control) are in fact due to area characteristics and largely independent of respondents' socio-demographic characteristics, which gives us some confidence that we are actually measuring environmental features when using small area units.

Second, we have demonstrated that the use of smaller area units produces more homogenous observations of the environment (as judged by the ICCs), indicating that the environment of smaller areas tends to be more homogeneous than that of larger areas. It should be stressed that when we talk about a homogenous environment in this context we could, in fact, refer to heterogeneous characteristics, such as ethnic diversity. For example, if all observers agree that an area is ethnically diverse their observations are homogeneous although this aspect of their environment is heterogeneous (but homogeneously heterogenic within its area boundaries). This is an important point because heterogeneity in environmental conditions is often of particular interest in the study of the role of the environment in crime causation (e.g., diversity in aspects of the area's population composition or land use). Thus when we refer to area homogeneity this includes (within-area) *homogeneous heterogeneity* in environmental conditions.

In our study, we compared area units averaging 300 inhabitants to area units averaging 1,500. Many ecological studies use area units averaging between 5,000 and 8,000 inhabitants (and some even more) and we would expect that the observed gain in area homogeneity using smaller area units would have been even more dramatic had our comparison involved such large units (although a city the size of Peterborough could not have been divided into enough units of this size to allow the ecometric analyses we have conducted).

Third, we have shown that an additional advantage of using more smaller area units rather than fewer larger units is an increase in statistical power for detecting significant area level effects. This result is in line with recent research on power analysis in multi-level survey designs showing that, in general, maximizing the number of area units is preferable to maximizing the number of respondents within each unit.

However, we have also demonstrated that using smaller areas with a lower number of respondents rather than fewer areas with more respondents inevitably affects reliability (as judged by the lambdas), which is a methodological disadvantage. In such case, the researcher must compromise between area homogeneity and reliability. This compromise will preferably be taken based on relevant theoretical and methodological concerns, such as aiming to approximate behavior-settings without losing too much in terms of reliability of data.

In this context it should, however, be stressed that there are no reasons, in principle, why the use of smaller area units should imply lower numbers of respondents. In practice, this is often the case generally because of financial constraints, which are a reality in almost all empirical research. It is more expensive, for example, to conduct a survey which has 30 respondents in each of 300 (small) areas than one which has 30 respondents in each of 100 (large) areas. However, we have shown that, on balance, the advantage of having more small areas (which are more homogenous) outweighs the loss of reliability.

So why is smaller better? Small units of analysis are better on *theoretical* grounds because they more closely approximate behavior-settings. Individuals' actions and development are only influenced by the environments they can access with their senses and the part of the environment which individuals can access with their senses is, arguably, generally small. Small units are also better from a *methodological* point of view because smaller units are more likely to be homogeneous in terms of environmental characteristics (although it is important to note that these homogenous environmental characteristics, in fact, can constitute heterogeneity, such as ethnic diversity or diversity in land use). Small units are also better because using more small areas provides more statistical power than using fewer large areas, making it easier to establish statistical significance. The only major drawback we have identified is that researchers may (purely for financial reasons) have to choose between the number of units and the number of respondents per unit which they include in a survey; opting for smaller units will mean they have fewer respondents per unit, which will affect the reliability of their estimates. By and large, however, in our evaluation, in order to advance the study of the role of the environment in crime causation small is certainly better.

## References


Bailey, T. C., & Gatrell, A. C. (1995). *Interactive spatial data analysis*. London: Longman.

Bellair, P. E. (1997). Social interaction and community crime: Examining the importance of neighbor networks. *Criminology*, 35, 677–703.

Hox, J. (2002). *Multilevel analysis: Techniques and applications*. Mawhaw: Erlbaum.

Kubrin, C. E., & Weitzer, R. (2003). New directions in social disorganization theory. *Journal of Research in Crime and Delinquency*, 40(4), 374–402.

Land, K. C., McCall, P. L., & Cohen, L. E. (1990). Structural covariates of homicide rates: Are there any invariances across time and space. *American Journal of Sociology*, 95(4), 922–963.

Martin, D. (2000). Towards the geographies of the 2001 UK census of population. *Transactions of the Institute of British Geographers*, 25, 321–332.

McCord, E. S., Ratcliffe, J. H., Garcia, R. M., & Taylor, R. B. (2007). Nonresidential crime attractors and generators elevate perceived neighborhood crime and incivilities. *Journal of Research in Crime and Delinquency*, 44(3), 295–320.

McVie, S., & Norris, P. (2006). *Neighbourhood effects on youth delinquency and drug use (Edinburgh study of youth transitions and crime, working paper 10)*. Edinburgh: Centre for Law and Society.

Miethe, T. D., & Meier, R. F. (1994). *Crime and its social context: Towards an integrated theory of offenders, victims, and situations*. Albany: State University of New York Press.

Murrey, D. M., Varnell, S. P., & Blitstein, J. L. (2004). Design and analysis of group-randomized trials: A review of recent methodological developments. *American Journal of Public Health*, 94(3), 423–432.

Oberwittler, D. (2003). Die Messung und Qualitätskontrolle kontextbezogener Befragungsdaten mithilfe der Mehrebenenanalyse - am Beispiel des Sozialkapitals von Stadtvierteln. *ZA-Informationen*, 53, 11–41.

Office of the Deputy Prime Minister. (2004). *The English Indices of Deprivation 2004: Summary*. London.

Openshaw, S., & Taylor, P. J. (1981). The modifiable areal unit problem. In: N. Wrigley & R. J. Bennet (Eds.), *Quantitative geography: A British view* (pp. 60–70). London: Routledge.

Openshaw, S. (1984). *Modifiable areal unit problem*. Norwich: GeoBooks.

Ouimet, M. (2000). Aggregation bias in ecological research: How social disorganization and criminal opportunities shape the spatial distribution of juvenile delinquency in Montreal. *Canadian Journal of Criminology*, 42, 135–156.

Raudenbush, S. W. (1999). Statistical analysis and optimal design in cluster randomized trials. *Psychological Methods*, 2(2), 173–185.

Raudenbush, S., & Bryk, A. (2002). *Hierarchical linear models: applications and data analysis methods (2.A.)*. Thousand Oaks: Sage.

Raudenbush, S. W., & Sampson, R. J. (1999). Ecometrics: Toward a science of assessing ecological settings, with appliance to the systematic social observation of neighborhoods. *Sociological Methodology*, 29, 1–41.

Reynolds, H. D. (1998). *The modifiable area unit problem: empirical analysis and statistical simulation*. PhD thesis, University of Toronto.

Sampson, R. J., Raudenbush, S. W., & Earls, F. J. (1997). Neighborhoods and violent crime: A multilevel study of collective efficacy. *Science*, 277, 918–924.

Sampson, R. J., & Raudenbush, S. W. (1999). Systematic social observation of public spaces: A new look at disorder in urban neighborhoods. *American Journal of Sociology*, 105(3), 603–651.

Sampson, R. J., Morenoff, J. D., & Gannon-Rowley, T. (2002). Assessing "neighborhood effects": Social processes and new directions in research. *Annual Review of Sociology*, 28, 443–478.

Sampson, R. (2006). How does community context matter? Social mechanisms and the explanation of crime rates. In: P.-O. H. Wikström & R. Sampson (Eds.), *Crime and its explanation: Contexts, mechanisms and development*. Cambridge: Cambridge University Press.

Shaw, C., & McKay, H. D. (1969 [1942]). *Juvenile Delinquency and Urban Areas*. Chicago: Chicago University Press.

Silver, E., & Miller, L. L. (2004). Sources of informal social control in Chicago neighborhoods. *Criminology*, 42(3), 551–583.

Smith, W. R., Frazee, S. G., & Davison, E. (2000). Furthering the integration of routine activity and social disorganization theories: Small units of analysis and the study of street robbery as a diffusion process. *Criminology*, 38(2), 489–523.

Snijders, T., & Bosker, R. (1999). *Multilevel Analysis: An introduction to basic and advanced multilevel analysis*. London: Sage.

Taylor, R. (2001). *Breaking away from broken windows: baltimore neighborhoods and the nationwide fight against crime, grime, fear, and decline*. Boulder, CO: Westview Press.

Taylor, R. (2002). Fear of crime, social ties, and collective efficacy: Maybe masquerading measurement, maybe Deja Vu all over again. *Justice Quarterly*, 19(4), 773–792.

Wikström, P.-O. H. (2004). Crime as alternative: Towards a cross-level situational action theory of crime causation. In: J. McCord (Ed.), *Beyond empiricism: Institutions and intentions in the study of crime* (pp. 1–37). New Brunswick: Transaction.

Wikström, P.-O. H. (2005). The social origins of pathways in crime. Towards a developmental ecological action theory of crime involvement and its changes. In: D. P. Farrington (Ed.), *Testing integrated developmental/life course theories of offending* (pp. 211–145). New Brunswick: Transaction.

Wikström, P.-O. H. (2006). Individuals, settings and acts of crime. situational mechanisms and the explanation of crime. In: P.-O. H. Wikström & R. J. Sampson (Eds.), *The explanation of crime: Context, mechanisms and development*. Cambridge: Cambridge University Press.

Wikström, P.-O. H. (2007a). The social ecology of crime: The role of the environment in crime causation. In: H.-J. Schneider (Ed.), *Internationales Handbuch der Kriminologie* (Vol. 1). Berlin: de Gruyter.

Wikström, P.-O. H. (2007b). Deterrence and deterrence experiences: Preventing crime through the threat of punishment. In: S. Shoham, O. Beck, & M. Kett (Eds.), *International handbook of penology and criminal justice* (pp. 345–378). London: CRC Press.

Wikström, P.-O. H. (2007c). In search of causes and explanations of crime. In: R. King & E. Wincup (Eds.), *Doing research on crime and justice* (2nd ed.). Oxford: Oxford University Press.

Wikström, P.-O. H., & Dolmen, L. (2001). Urbanisation, neighbourhood social integration, informal social control, minor social disorder, and fear of crime. *International Review of Victimology*, 8, 121–140.

Wikström, P.-O. H., & Ceccato, V. (2004). *Crime and social life: A space-time budget study*. Paper presented at the annual meeting of the American Society of Criminology. Nashville, TN. November 2004.

Wikström, P.-O. H., & Butterworth, D. (2006). *Adolescent crime: Individual differences and lifestyles.* Collumpton: Willan Publishing.

Wikström, P.-O. H., & Treiber, K. (2007). The role of self-control in crime causation: Beyond Gottfredson and Hirschi's general theory of crime. *European Journal of Criminology*, 4(2), 237–264.

Wikström, P.-O. H., & Treiber, K. (2008). What drives persistent offending? The neglected and unexplored role of the social environment. In: J. Savage (Ed.), *The development of persistent criminality*. Oxford: Oxford University Press.

Wooldredge, J. (2002). Examining the (Ir)Relevance of aggregation bias for multilevel studies of neighborhoods and crime with an example comparing census tracts to official neighborhoods in Cincinnati. *Criminology*, 40(3), 681–709.

http://www.springer.com/978-0-387-09687-2

Putting Crime in its Place
Units of Analysis in Geographic Criminology
Weisburd, D.; Bernasco, W.; Bruinsma, G. (Eds.)
2009, XII, 254 p. 55 illus., Hardcover
ISBN: 978-0-387-09687-2