Eyes on the Block: Measuring Urban Physical Disorder Through In-Person Observation

Malia Jones
UCLA School of Public Health

Anne R. Pebley
UCLA School of Public Health

and

Narayan Sastry
University of Michigan, Institute for Social Research

April 2010

The authors gratefully acknowledge support from NICHD (R01 HD35944 and R01 HD41486) and from the Russell Sage Foundation's Social Inequality Project.

Introduction

A large literature suggests that physical disorder (e.g., dilapidated buildings, trash, broken sidewalks, graffiti, etc.) in urban areas may increase the crime rates, disruptive behavior, stress levels, and health and psychological problems among neighborhood residents (Hill et al., 2005; Ross and Mirowsky, 2001; Taylor, 2001). The theory that reducing physical disorder can increase safety and social control has led to anti-graffiti programs, vacant land management, and other initiatives to reduce "urban decay" in many cities (Barnard, 2006; Wachter and Gillen, 2006).[1] However, researchers have only recently begun to develop methods of measuring urban physical disorder (Raudenbush and Sampson, 1999; Taylor et al., 1985). Comprehensive and reliable measures are essential to testing hypotheses about the consequences of physical disorder and to identifying urban communities where disorder is most acute.

In this paper, we examine the results from the Los Angeles Family and Neighborhood Survey, Wave 1 (L.A.FANS-1). In L.A.FANS-1, multiple trained observers walked through neighborhoods while systematically recording what they observed on standardized forms. Although expensive, independent neighborhood observations can provide more objective assessments of neighborhood conditions than proxy measures used by most studies.[2] We examine this method of assessing physical disorder in Los Angeles by addressing three questions. First, do independent observers of the same neighborhood perceive the same levels of physical disorder? If independent

[1] The "broken windows" theory of the effects of physical and social disorder on crime rates has also led, more famously, to changes in urban policing strategies focused on reducing petty crime (Bratton and Kelling, 2006). Although policing strategies are an important policy issue, they are outside the scope of this paper.
[2] One limitation of independent observations is that they do not capture residents' perceptions or feelings about specific aspects of physical disorder (Raudenbush and Sampson, 1999), but this information can be collected in other ways, e.g., through interviews with residents themselves.

observers agree about disorder levels, the items that they were asked to code are more likely to be well-defined, readily observable, and replicable. Second, how do circumstances (time of day, day of the week and season) of observation and observers' prior experience with the neighborhood affect their perceptions of physical disorder? Earlier research and common sense suggests that what is observed in a neighborhood may be different in the morning and evening (Raudenbush and Sampson, 1999), on weekdays and weekends, and in different seasons—although the effect of situational variables may be quite modest for items that change little over time (e.g., the number of street lanes) compared to items for which there may be substantial variation over time (e.g., presence of litter). Observers' previous experience in the neighborhood may also affect what they observe, despite standardized training and methods. We assess how important these factors are in determining what observers report. The third question we address is what types of neighborhoods in Los Angeles County have the highest levels of physical disorder? In answering this question, we also examine how well social characteristics available from sources such as the U.S. Census serve as proxies for more comprehensive and direct measures of neighborhood disorder.

Our results indicate that there is considerable variation in the level of agreement among independent observations across items that capture different aspects of neighborhood physical disorder (e.g., litter, graffiti, well-tended yards). Overall, levels of agreement among different observers of the same block face are moderate to high. The day of the week and season in which the observation was made have statistically significantly effects on the observers' perceived level of disorder, as do the observers' previous experiences in the neighborhood. However, the effect of these situational variables on the degree of agreement is modest. There is substantial variation in the correlation of block face observations by item, with results largely reflecting the degree to which there are systematic differences in physical characteristics across block faces in the same tract. Characteristics that are likely to be shared across all block faces in a tract—such as the presence of graffiti—have substantially less inter-block face variability than those that vary substantially by block face—such as the number of traffic lanes. The former items are likely to be more useful for constructing tract-level summary scales of physical disorder. Our results also show that physical disorder is significantly related to neighborhood poverty, affluence, residential stability and race/ethnic diversity, but not to immigrant concentration. These characteristics combined explain much of the tract level variation in physical disorder.

The Effects of Physical Disorder

Physical disorder includes the condition of streets, sidewalks, building exteriors, and other characteristics visible to any passer-by. Both the criminology and social epidemiology literatures have been concerned with the consequences of urban physical disorder for neighborhood and individual welfare. Below, we provide a brief overview of each literature.

<u>Physical Disorder and Crime</u>

The "broken windows" theory first advanced in the 1980s (Wilson and Kelling, 1982) suggests that physical disorder affects crime rates, in two ways. First, physical disorder including broken windows and graffiti visually advertises that residents tolerate infractions against social order and are unlikely to intervene to stop crime and disorderly conduct. To potential criminals, disorder indicates poor social control, which they can exploit. Second, residents feel personally threatened by disorderly elements such as drunks, graffiti, and trash. Therefore, they retreat into their homes, spending less time in public spaces and investing less in relationships with neighbors (Skogan, 1990; Wilson and Kelling, 1982). As a consequence, few people spend time on the street, undermining social control and further increasing opportunity for disorderly conduct and crime. Jane Jacobs believed this dearth of "eyes on the street" was a key mechanism of urban decay (Jacobs, 1961; Skogan, 1990). The neighborhood's ability to act collectively (e.g., to maintain order) is also impaired because residents do not know or trust each other (Sampson and Raudenbush, 1999). Residents are less likely to intervene to prevent disorderly conduct and crime because they fear that their neighbors will not back them up or may even threaten them.

Physical disorder can generate further disorder, since it encourages residents to move to less disordered neighborhoods (Skogan, 1990). Because this option is less available to poor residents, disorder can contribute to a concentration of poverty and to disinvestment in housing and businesses in the area (Wilson, 1987). The remaining residents are less enfranchised, have fewer resources, and may feel less ownership of their streets (Sampson and Raudenbush, 1999; Skogan, 1990). Because they are disenfranchised, they are also less likely to maintain public spaces and to keep physical disorder at bay (e.g., picking up litter, painting over graffiti, and maintaining yards and buildings).

Cross-sectional studies show that neighborhood disorder is associated with higher crime rates and fear of crime (Kelling and Coles, 1998; Sampson and Raudenbush, 1999; Skogan, 1990). However, crime may, in fact, cause perceived disorder rather than the reverse (Harcourt, 2001). Sampson and Raudenbush (1999) argue that disorder is a symptom, not a cause, of poor social control and crime in the neighborhood. Therefore, cleaning up minor infractions such as graffiti and cigarette butts will have little effect on burglaries and homicides. Rather, the solution to both problems is to improve social control through strengthening trust and collective efficacy within neighborhoods.

<u>Physical Disorder and Stress</u>

Research in social epidemiology suggests that physical disorder can also cause chronic stress among neighborhood residents. Residents may experience fear of crime and violence, feelings of hopelessness, or feelings of isolation, each of which is a source of chronic stress. Chronic stress, in turn, increases the risk of negative health outcomes such as obesity, high blood pressure, heart disease, and depression (Hill et al., 2005; McEwen, 1998; Molnar et al., 2004; Ross and Mirowsky, 2001; Sampson and Raudenbush, 1999). Reducing physical disorder and chaotic living conditions may, therefore, improve the mental and physical health of neighborhood residents—particularly in poor communities, where both physical disorder and poor health are more common.

Empirical evidence on the disorder-health relationship is limited and often indirect. For example, Moving to Opportunity (MTO) – a randomized intervention in which selected residents of subsidized housing were moved to middle class neighborhoods – found that mental health and feelings of safety improved significantly among adults and female youth in the treatment group compared to control group members who remained in poor neighborhoods (Kling et al., 2007). However, this study did not directly examine physical disorder.

Ross and Mirowsky (2001) found that residents reporting high neighborhood physical disorder also reported worse health status and more limitations in physical functioning than their counterparts in neighborhoods with less disorder. Their models suggest that stress, specifically fear of crime and violence, is the mechanism linking health to physical disorder. Their study relies on self-reported disorder as well as self-reported fear and health, which raises some questions about the causal order of variables.

Neighborhood Social Characteristics and Physical Disorder

Socially disadvantaged neighborhoods appear to have higher levels of physical disorder (Cohen et al., 2003; Sampson and Raudenbush, 1999), for several reasons. Residents of poor neighborhoods have less income, and often less time, to maintain their homes, yards, public spaces, and businesses. Less political clout makes it harder to obtain public maintenance services such as graffiti cleanup and sidewalk repair. Businesses in poor neighborhoods have also fewer resources to maintain their property and contribute to neighborhood improvement projects (Alwitt and Donley, 1997). Under-investment can lead to vacant or abandoned properties, which are themselves a form of physical disorder (Sampson and Raudenbush, 1999).

High levels of residential turnover and a high proportion of renters may also contribute to physical disorder. Frequent residential turnover makes it difficult for neighbors to get to know each other, establish trust, and exercise social control (Ross and

Jang, 2000; Sampson et al., 1997). High homeownership rates are associated with residential stability and may improve both property maintenance and residents' ability to control their physical surroundings (Sampson and Raudenbush, 1999). Homeowners have more incentive to invest in home and neighborhood improvement than renters because it improves property values. For the same reasons, they are also more likely to share with other homeowners norms about appropriate behavior (e.g., trash disposal, building and yard maintenance).

Land use patterns may also affect the ability of residents to exercise social control. The land on which poor neighborhoods are built is more likely to be located in or near industrial sites or other "waste lands," or close to urban infrastructure (e.g., freeways, bridges, train tracks). These "no man's lands" are not controlled by the land's owners or by local residents. Both outsiders and residents feel free to dump trash or other waste in these areas. They can also be havens for graffiti, illegal camping, and crime. Physical and social disorder from these areas can spill over into residential areas of poor neighborhoods (Mohl, 2002).

Immigrant concentration may also contribute to physical disorder, independently of poverty and residential instability. Where immigrant neighborhoods are more culturally and linguistically diverse, residents may be less likely to form social bonds, and therefore to collaborate to limit or remove physical disorder (Sampson et al., 1997). In Los Angeles, the immigrant population comprised more than one third of population in 2000 (Social Science Data Analysis Network (SSDAN), No date). However, immigrant neighborhoods in Los Angeles are, on average, *less* ethnically diverse than other neighborhoods: the high level of immigration from Mexico and Central America combined with residential settlement patterns has generated highly concentrated immigrant neighborhoods. In 2000, the average Latino person in Los Angeles lived in a census tract that was 78% Latino (Ortiz and Telles, 2008). Although less ethnically diverse, immigrant neighborhoods in Los Angeles are often more disenfranchised and less able to marshal resources or obtain public services. For this reason alone, they may have higher levels of physical disorder.

Thus, previous research suggests that physical disorder is likely to be more common in neighborhoods that are poor, have high residential turnover rates, have lower owner-occupancy rates, are near to industrial sites or freeways, and have high concentrations of immigrants.

Measurement of Physical Disorder

Although "windshield surveys" have a long history in urban studies and public health, most research on neighborhood disorder relies on residents' perceptions of disorder (Perkins and Taylor, 1996; Ross and Mirowsky, 2001; Sampson and Raudenbush, 1999). Several recent studies have attempted to develop independent and objective measures of disorder. In an early study, Perkins et al. (1992) focused on observation of three theoretical constructs in Baltimore neighborhoods: physical and social incivilities (e.g., litter, vandalism, harassment, selling drugs), territorial functioning (e.g., property maintenance, "neighborhood watch" signs), and defensible space (e.g., lighting, fences). Each sampled block was observed simultaneously by two trained observers who were instructed not to discuss ratings with each other. Intra-class correlation coefficients (ICCs) and Cronbach's alpha statistics calculated for the two observers' rating for each block were remarkably high for many physical characteristics. For example, the ICCs for the estimated percentage of open block frontage that was unused vacant lots, parking lots, public playgrounds, and public gardens ranged from 0.97 to 0.99, suggesting either that the two observers were highly consistent or, possibly, that the observers at least sometimes violated the prohibition on discussing observations. However, agreement on some items (e.g., number of abandoned cars on the street) was considerably lower (Perkins et al., 1992): Table 1).

In the Project on Human Development in Chicago Neighborhoods (PHDCN), Sampson and colleagues (Raudenbush and Sampson, 1999; Sampson and Raudenbush, 1999) used a motor vehicle with videotape cameras on each side and one observer for each side of the block to observe social disorder (e.g., adults loitering, people drinking) and physical disorder (e.g., litter, graffiti). Videotapes were coded by independent observers and differences reconciled. Several other studies have used shortened or modified versions of the PHDCN instruments(Franzini et al., 2009; Grafova, 2008; Kelly et al., 2007; Wei et al., 2005) while some studies have developed their own questions on physical disorder (Miles, 2006). The most common items in all these studies include the presence or absence of litter and garbage, graffiti, beer and liquor containers, broken glass, abandoned cars, vacant lots, condoms, and drug paraphernalia.

Most of these studies did not assess inter-rater reliability or the effects of observer characteristics on the observation results. For example, in PHDCN, only one person in the vehicle observed each block face. Although several coders coded the videotapes independently, the authors do not report inter-rater agreement among them. However, when a random 10 percent of all block faces were recoded from the videotapes by new coders, the level of agreement between the old and new coders was 98 percent (Raudenbush and Sampson, 1999).

In Wei et al. (2005), only one observer completed the observations, but in 5% of the blocks, a researcher subsequently reassessed the block. The ICC for the two observations was 0.68. In several studies, only one person observed a given block face (Grafova, 2008; Miles, 2006) and in other studies where two or more people completed the observation it is not clear whether they conferred about the results (Franzini et al., 2009; Kelly et al., 2007).

Data and Measurement

We use data from Wave 1 of the Los Angeles Family and Neighborhood Survey (L.A.FANS-1) conducted in 2000-2001 in Los Angeles County. To assess physical disorder, L.A.FANS-1 trained observers to record levels of disorder using a standardized list of items (e.g., presence of cigarette butts, bottles and cans, building condition, etc.). Each block face was observed by multiple observers working independently. Furthermore, all observers conducted observations in multiple tracts and block faces. The specific tracts and blocks that observers were assigned to visit was based on the observers' residential location in order to reduce travel costs (because they worked from home); this procedure also resulted in a correlation between observers' race and ethnicity and that of the neighborhoods they observed. The non-random assignment of observers to tracts means that attempting to control for correlation in observations performed by the same observer is problematic because it may yield spurious findings. The scheduling of observations and the time to completion were also variable.[3] The observation forms were adapted from those used in the Project on Human Development in Chicago Neighborhood (PHDCN).

L.A.FANS is based on a stratified probability sample of census tracts (defined by 1990 boundaries). Three strata were defined based on the percent in poverty in 1997: very poor (in the top 10 percent of the poverty distribution), poor (in the next 30 percent), and non-poor (in the bottom 60 percent). Tracts in the very poor and poor strata were oversampled. In each tract, census blocks were sampled with probability proportional to population size. An average of 6.6 blocks was selected per tract, with a range of 2 to 14 blocks (Sastry et al., 2006). Observations were completed by specially trained L.A.FANS interviewers for each block face in sampled blocks. Observers first drove around the entire block, then walked along each block face, observing both sides of the street. At the end of each block face, the observer completed a standardized form. Signs of both physical and social disorder were recorded, but the social disorder indicators are not included in this analysis because of low frequency and poor reliability.

Basic characteristics of the block face sample are shown in Table 1. In total, the L.A.FANS sample includes 2,071 block faces, 422 blocks, and 65 tracts. On average, each block comprised of about 5 block faces and there were 6.5 blocks per tract. An average of about 3 independent observations was completed per block face, 14 per block, and 92 per tract. Observers completed a mean of 171 observations. For 98 percent of block faces, more than one observation was completed and three or more observations were completed for 80 percent of block faces.

[3] Under optimal conditions both the observers and the time of visits would be randomly assigned across block faces. However, field conditions and budget prevent this randomization in L.A.FANS.

Table 2 shows summary statistics for the situational characteristics of the observations. Observations were fairly evenly spread across the days of the week although they were less likely to be completed on Mondays. About three-quarters of observations were completed at midday or in the afternoon. Almost all observations were completed in the winter. The reason for the strong seasonal clustering was that most observations were completed during a single fieldwork campaign which occurred in the winter months – when the temperature is generally cool, but not cold, and there is very occasional rain. Most observations were completed by observers who had no previous knowledge of the block being observed. The average time spend observing a block face and completing the form was one hour, although it is important to note that there was considerable variation among observations in this length as indicated by the standard deviation.

The block face observation form contained 40 items.[4] Specific items are shown in Table 2, along with their minimum and maximum values, mean, and standard deviation. The column labeled "type" indicates whether the variable is continuous (C), ordered (O), or dichotomous (D). Most responses were recorded using one of two Likert-style response scales. One scale contained four response options: "none", "a little", "some", or "a lot". The other contained five response options: "none", "very few", "some", "many", or "all". These variables are categorized as ordered (O). A few items were recorded on the original form as "yes" or "no" for the presence or absence of a condition. These are listed as dichotomized variables (D).

[4] The observation form and manual are available at: www.rand.org/pubs/drafts/2005/DRU2400.6-1.pdf

Sampson and Raudenbush found that most of the systematic variation in neighborhood observations in PHDCN was captured by dichotomizing the responses into "none" and "any" categories (Raudenbush and Sampson, 1999; Sampson and Raudenbush, 1999). To determine whether this finding is equally true in L.A.FANS-1, we conducted two analyses. First, we compared, at the block-face level, ICCs for each variable, dichotomized at each possible cutoff point in the scale (results not shown). The ICCs were almost always highest when data were dichotomized between "none" versus "any" categories, suggesting that this scheme was the most appropriate. Second, we conduct the analysis of inter-rater reliability for both the ordered form and the dichotomized form of these variables, to determine whether the results differ depending on specification, as described below.

The strongly residential and suburban character of L.A.FANS blocks (and most of Los Angeles County) is apparent from the traffic, land use, and building type variables. Streets were generally narrow residential streets with two lanes of traffic. The primary land use was residential and the majority of block faces contained stand-alone houses – the predominant housing type in Los Angeles County. Although observers coded 10 percent of block faces as having no residential housing, all blocks selected for L.A.FANS had some residential housing according to the 1990 census and other administrative data (Sastry et al., 2006).

To assess the association of physical disorder with neighborhood economic, social, and demographic characteristics, we used five tract-level summary variables created by the L.A.FANS project for all Los Angeles tracts (Peterson et al., 2007) using data from the 2000 U.S. Census.[5] Data from the 2000 Census were matched with 1990 tracts using a standard cross-walk. The five variables are concentrated disadvantage, concentrated affluence, ethnic diversity, residential stability, and immigrant concentration indices. The concentrated disadvantage index includes the percent of the population: in poverty, with annual family income <$24,000, in female headed households, receiving public assistance, nonwhite, and <18 years old. The concentrated affluence index includes the percent: in executive or professional occupations, with 13+ years of schooling (adults 25+ only), with annual family income >$75,000, white, who speak English "very well" (among adults), and who speak only English (among adults). The ethnic diversity index reflects the probability that any two people chosen at random from the tract would be of different race/ethnicity. For this measure, race/ethnic groups are defined as: Latino, white, African American, Asian, and other. The score is calculated as:

$$1 - ([\%\ white]^2 + [\%\ black]^2 + [\%\ Latino]^2 + [\%\ API]^2 + [\%\ other\ race/eth]^2) \quad (1)$$

The residential stability index was constructed using factor analysis and includes percent of: dwellings in multi-unit housing, owner-occupied housing units, households living in the same residence as 5 years earlier, and non-family households. The immigrant concentration index, also constructed using factor analysis, includes the percent of the population: non-citizen, foreign born, foreign born who arrived since 1990, foreign born who arrived since 1995, Spanish speaking, and Latino. In initial multivariate models (not shown), we also included population density and land use.

---

[5] This analysis is done at the tract level because census data on social characteristics are not available at the block level. They are available at the block group level, but since L.A.FANS observations were conducted only in sampled blocks, for many block groups, not all blocks in the group were observed.

However, the coefficients for these variables were not statistically significant and are omitted from analyses presented here.

Analytic Approach

Our first two goals are to determine the extent to which independent observers of the same block face recorded the same levels of physical disorder – i.e., the degree of inter-rater reliability – and to investigate whether situational characteristics (time of day, day of the week and season) affected perceptions of physical disorder. We also assess whether dichotomized versions of the ordered observation items perform as well as the original ordered items themselves on inter-rater reliability.

We use ICCs to measure inter-rater agreement. Kappa statistics are sometimes used for this purpose, but they cannot be used to compare the amount of agreement across variables (Landis JR and Koch GG, 1977). ICCs, on the other hand, provide a continuous measure of correlation that represents the degree of agreement among independent observers. The magnitude of the ICC provides a means of comparing the strength of agreement on one aspect of disorder (e.g., trash) vs. other aspects (e.g., conditions of buildings). We estimate ICCs using a set of multilevel models with random effects at the block face, tract, or both block face and tract. We also estimate models with and without the situational variables to assess how these variables affect the reliability of observation among observers.

Although examining each observed characteristic can be informative, our central interest is in the latent or composite construct of neighborhood physical disorder. We created two summary scales to measure this construct. The first scale approximates

Sampson and Raudenbush's (1999) physical disorder scale from PHDCN data. The PHDCN scale included 10 items: cigarettes/cigars, garbage or litter, empty beer bottles, three types of graffiti (political, tagging and gang), painted-over graffiti, abandoned cars, condoms, and needles/syringes. In the L.A.FANS observation form, the three types of graffiti were combined into one item.[6] We also combined the presence of condoms and needles/syringes into a single item, because both were rarely observed. Thus, the L.A.FANS scale is based on 7 rather than 10 items.

The PHDCN scale is based on a relatively narrow definition of physical disorder. Therefore, we created a second scale using factor analysis with varimax rotation to select the items to be included. The analysis included all observed physical characteristics of the block face. Three or four strong factors emerged from this analysis, depending on the model constraints (results not shown). Only one of the factors (the strongest) has a clear theoretical interpretation related to poor maintenance, safety-related devices, and multi-unit dwellings. The twenty items most strongly associated with the first factor were selected. They include the items in the PHDCN scale. This 20-item scale has a Cronbach's α of 0.90, and a slightly right-skew.

Next, we estimated multilevel item response models to analyze the neighborhood physical disorder scales. These models also allow us to assess the role of situational variables (time of day, season, etc.) in affecting the reliability of reporting among observers of the same place. We estimated standard one-parameter item response models with random effects at the block face and tract levels (Raudenbush and Sampson, 1999). The dependent variables in these models are the binary item responses for the entire set

[6] The L.A.FANS item on the presence of beer bottles was also broader than in the PHDCN: referring to both "beer containers and liquor bottles."

of 7 or 20 items (for each of the two scales). The models were estimated using multilevel logistic regression via maximum likelihood estimation procedures. These models can be viewed as three-level models, with the first level being item responses within the block face, the second level being block faces, and the third level being tracts. Covariates that reflect situational characteristics appear at the first level of the model, along with dummy variables that identify each individual item (the models do not include a constant). The coefficients on the item dummies represent estimates of item severity, in terms of its contribution to tract-level disorder. The model-based estimate of the random effect for each tract represents the physical disorder scale for that tract. We also calculated tract-level ICCs for the scales using the estimated variances of random effects at the block face and tract levels.

In the final part of the analysis, we used the tract-level physical disorder scale from the two item response models as dependent variables in a linear regression analysis of the association between tract-level structural factors and physical disorder. Variables describing the tract structural characteristics are derived from 2000 Census data, as described above. Our goal is to assess which types of tracts are most likely to have high levels of physical disorder and how effectively census tract characteristics serve as proxy variables for direct observation of neighborhood physical disorder in this setting.

Results

Observer Agreement

Table 4 shows the ICCs from the multilevel random effects models without and with the situational variables. Model 1 includes a block face-level random effect and no covariates while Model 2 adds the situational variables. Similarly, Models 3 and 4 include tract-level random effects, without and with situational variables, respectively. Model 5 includes both block face-level and tract-level random effects in the same model plus the situational variables.

There are three key results in Table 4. First, the level of inter-rater agreement at the block face level is generally high, ranging from a low of 0.32 to a high of 0.95 (Model 1). In Models 3 and 4, the ICCs at the tract level are generally lower than at the block face level (Models 1 and 2), although the overall level of inter-rater agreement at the tract level is moderate to high (range for Model 3: 0.11 to 0.76). However, in Models 1 to 4, block face and tract effects are not clearly distinguished because the included effect picks up the influence of the omitted effect. Model 5 includes both tract and block face random effects simultaneously and hence provides net estimates of the block face and tract ICCs. In contrast to the results in the preceding columns, the ICCs estimated in Model 5 are sometimes considerably larger at the tract level than the block face level and sometimes the reverse. Not surprisingly, block face ICCs are generally larger than tract ICCs for items which are block face-specific and exogenously generated such as the number of traffic lanes, availability of public transportation, and whether there are street barricades and trees. By contrast, tract-level ICCs are higher for items reflecting conditions that are common throughout the entire local area, such as graffiti, abandoned cars, and trash and garbage. Note that in Model 5, the tract ICCs are larger than the block face ICCs for the variables that are included in the 7-item scale (items 10 and 11–17), described below.

The second key result in Table 4 is that holding constant the situational variables makes little difference in the size of the ICCs (e.g., contrast model 1 with model 2 and model 3 with model 4). This finding suggests that relatively little of the divergence in ratings among observers is due to variations in the circumstances of the observations such as season of the year or day of the week.

Third, for the ordered variables based on Likert-type scales, dichotomizing them into "none" and "any" results in only minor changes to the ICCs in all of the five model specifications. The ICCs for the dichotomized variables are often very close to, and occasionally larger than, the ICCs for the ordered versions of the variables. This finding is consistent with Raudenbush and Sampson's (1999) results from PHDCN and suggests that it is difficult for observers to distinguish consistently between qualitative descriptions such as "a little," "some," and "a lot."

Multilevel Item Response Models

The results in Table 4 suggest that situational variables account for very little of the *variation among observers* in their observations. Next we want to assess how much the situational variables affect the likelihood of observing physical disorder, across items, observers, block faces and tracts. We also want to assess the degree of severity for each item in each of the scales. To do so, we estimate multilevel item response models. The parameter estimates and standard errors from these models for the 7-item PHDCN-equivalent scale and the 20-item expanded scale are shown in Table 5.

At the top of the Table 5 are estimated coefficients for each of the individual observation items that comprise the scale. Items with lower probabilities of occurrence (see Table 3) have more negative coefficients, while those with higher probabilities of occurrence have larger positive coefficients. For the 7-item scale, the item with the lowest probability of occurrence—observed drug paraphernalia, observed in just 3 percent of block faces—has the most negative coefficient, while the item with the highest probability of occurrence—the presence of garbage, litter or glass, observed in 73 percent of block faces—has the largest positive coefficient. A similar pattern of findings holds for the 20-item scale. The coefficients thus correspond to the degree of severity of the observation item, in terms of its contribution to neighborhood disorder. Graffiti is the only item in the 7-item scale that does not have an estimated coefficient significantly different from zero. This result is related to the fact that graffiti was observed in approximately half of all block faces, and hence this item is unlikely to be strongly related to whether a tract is high or low on the neighborhood disorder scale. A parallel set of findings emerge for the 20-item scale. The large observed variation in the estimates of item severity is an indication that the scale is well-behaved (Raudenbush and Sampson, 1999).

The situational variables are shown next in Table 5. The estimated parameters are interpreted as systematic effects—across items, block faces, tracts, and interviewers—on the likelihood of observing disorder for the corresponding situational variable. These variables are jointly significant in both the 7-item and 20-item models. With only one exception—the effects of time of day in the 20-item model—all of the variables are statistically significant, based on a set of joint tests for the multiple categorical variables used to characterize each discrete variable (results not shown). The magnitudes of the estimated parameter effects are, however, relatively modest, with the largest covariate effects found for season. Nevertheless, the results suggest that the likelihood of observing disorder decreased with the duration of time spent observing; was highest if the observation was conducted mid-week (Wednesday) or on the weekend (either Saturday or Sunday); was higher if the observation was conducted around midday (only for the 7-item scale); was higher in the summer and winter; and was higher if the interviewer knew the neighborhood through the L.A.FANS study.

At the bottom of Table 5, we show the variance of the block-face and tract random effects. Both random effects are statistically significant, with the variance of tract random effect about three times larger than the variance of the block-face random effect for both the 7-item and the 20-item models. In Table 6, we present the corresponding ICCs for block face and tract. Each of the ICCs is larger for the 7-item scale. The tract ICC is 0.36 for the 7-item scale and 0.30 for the 20-item scale. The tract ICCs are about three-times larger than the block-face ICCs, which are 0.14 for the 7-item scale and 0.10 for the 20-item scale. The ICCs are similar to those for each of the individual items (shown in Model 5 in Table 4).

In the bottom panel of Table 6 we present summary statistics for the tract physical disorder scales that were obtained from the models presented in Table 5 as predicted values of the tract-specific random effects. The 7-item scale has a mean of zero, a standard deviation of 1.54 and a range of –2.71 to 2.35. The scale values have the same metric as the estimated coefficients for the individual observation items. Thus, as Raudenbush and Sampson (1999) point out, differences between tracts in their scores on the physical disorder scale can be interpreted as expected differences in the log-odds of finding disorder across the items in the scale. The interpretability of the scale and its well-behaved distributional properties mean that we can use it to characterize neighborhood physical disorder and to analyze the causes and consequences of disorder.

Neighborhood Social Characteristics and Physical Disorder

Finally, we examine what types of tracts in Los Angeles County have the highest levels of physical disorder. Previous research suggests that high poverty, ethnically diverse, residentially unstable, and high immigrant neighborhoods are more likely to have significant physical disorder. We test these hypotheses using tract-level multivariate models in which the 7-item scale and the 20-item scale are regressed on variables (discussed above) that describe tract levels of concentrated disadvantage, concentrated affluence, immigrant concentration, residential stability, and ethnic diversity.[7]

The results are shown in Table 7. The results for the 7-item scale and the 20-item scale are similar and the two models fit the data equally well, based on the model F-statistic reported at the bottom of the table. Physical disorder is significantly higher in disadvantaged neighborhoods and lower in affluent ones. Although concentrated disadvantage and concentrated affluence are highly correlated with each other (r = –0.88), initial analysis (not shown) revealed that together they perform better in predicting neighborhood disorder than either does on its own. These two variables also have a large combined effect: together they explain about 84% of the variation in disorder with no other variables included in the model (results not shown). The strong predictive value of these two variables suggests that they would be effective proxies for physical disorder, at

[7] Initial analysis (not shown) compared models estimating the effects of concentrated affluence and disadvantage variables on the 20-item scale with models including other specifications of neighborhood disadvantage (i.e., concentrated disadvantage alone, concentrated affluence alone, the percent of the population in poverty, median family income, percent of families early more than $75,000 per year, percent receiving public assistant, and percent female-headed households). The combination of concentrated disadvantage and affluence indices accounted for the greatest variance (adjusted R^2) and had the lowest Bayesian Information Coefficient (BIC) across models.

least in Los Angeles County in 2000-2001. This result contrasts to the findings of Raudenbush and Sampson (1999) who show that the correlation between concentrated poverty and physical disorder in Chicago neighborhoods, based on data from the PHDCN, was about 0.64—substantially lower than what we find for Los Angeles neighborhoods.

Contrary to our expectation, immigrant concentration is not significantly associated with physical disorder. Ethnic diversity is associated with a statistically significant *decrease* in disorder (marginally significant in the model based on the 7-item scale), contradicting the hypothesis that residents in diverse neighborhoods find it more difficult to exercise social control over physical disorder. This finding may reflect race/ethnic segregation patterns in Los Angeles because ethnically homogeneous tracts are predominately poor and Latino. Thus, in this setting, diversity may serve as another indicator of more advantaged neighborhood status. Greater residential stability is associated with lower levels of physical disorder, supporting the hypothesis that higher neighborhood turnover rates make it more difficult for residents to exert control over their neighborhood environment.

Conclusion

Concern about the potentially pernicious effects of physical and social disorder on residents of poor urban neighborhoods has pervaded policy and academic discussion in many fields. However, the development of reliable measures of disorder has lagged until recently. In this paper, we evaluated a strategy for assessing physical disorder conducted on foot by multiple trained independent observers in a stratified probability sample of neighborhoods in Los Angeles County.

Unlike many previous studies, the L.A.FANS data allow us to investigate the reliability of measures of specific aspects of physical disorder and the effects of interviewer and situational variables. Our results show that inter-rater agreement levels are generally high for multiple observers of the same block face and that these levels vary considerably by the item observed. More subjective and transitory aspects of disorder – e.g., garbage, strong odors, drug paraphernalia, cigarette butts, and beer or liquor containers – have lower levels of agreement than more enduring and objective aspects – e.g., lanes of traffic, vacant lots and abandoned buildings. This is an important finding because observational measures used in studies often include or are limited to the more subjective and transitory items to measure physical disorder. These results suggest that less ephemeral indicators of disorder such as vacant lots and abandoned buildings may provide a more reliable measure of neighborhood conditions – although these types of disorder may be less within residents' power to control.

Levels of disorder observed were modestly affected by the length of time the observation took, day of the week, time of day, season, and the observers' previous experience in the neighborhood. These results suggest that fieldwork designed to assess physical disorder should seek to minimize variation in scheduling neighborhood observations across days of the week, time of day, and season or, more realistically, should control for these variables in models based on neighborhood observations. Whenever possible, studies should also employ multiple trained observers to code each block face to assess inter-observer agreement and the effects of observers' characteristics on the level of disorder observed. Multiple observations of each location also allow studies to improve the quality of observations by creating variables which remove the effects of interviewer characteristics as we have done in this paper. Increasing the number of independent observations also improves the reliability and the precision of the estimated neighborhood physical disorder scales.

In Los Angeles County, concentrated disadvantage and affluence combined are strong predictors of physical disorder. Residential stability is significantly associated with lower physical disorder. Contrary to expectation, higher levels of ethnic diversity are associated with less, rather than more, physical disorder. The reason may be that neighborhoods with low ethnic diversity in Los Angeles are predominantly Latino neighborhoods, which are more likely to be disadvantaged in other ways. In contrast to Sampson and Raudenbush's (1999) findings that Chicago neighborhood with high levels of immigrant concentration had significantly more physical disorder, in our study, the coefficients on immigrant concentration were not significant – suggesting a very different effect of immigrant characteristics and settlement in the two cities.

Neighborhood observation by trained observers is an important means of measuring physical disorder in large social surveys. Our results indicate the importance of high quality training of observers, consideration of which aspects of physical disorder are more reliably observed, and, when possible, the use of multiple independent observers to allow researchers to examine the reliability of observations and improve the quality of the derived neighborhood scales.

In future work, we plan to assess the properties of neighborhood physical disorder scales constructed at the block level using the L.A.FANS observation data. We will also use both the block and tract physical disorder scales as explanatory variables in models of health and well-being outcomes assessed among individual L.A.FANS respondents. Our goal for this planned analysis is to determine the extent to which physical disorder is an important pathway through which neighborhood structural factors—such as concentrated disadvantage and affluence—influence outcomes such as obesity, depression, and children's behavioral problems.

Bibliography

Alwitt L.F., Donley T.D. (1997) Retail Stores in Poor Urban Neighborhoods. Journal of Consumer Affairs 31:139-164.

Barnard L. (2006) Graffiti Abatement and Management. Law and Order 50:115-119.

Cohen D.A., Farley T.A., Mason K. (2003) Why is poverty unhealthy? - Social and physical mediators. Social Science & Medicine 57:1631-1641.

Franzini L., Elliott M.N., Cuccaro P.*, et al.* (2009) Influences of physical and social neighborhood environments on children's physical activity and obesity. Am J Public Health 99:271-8.

Grafova I.B. (2008) Overweight children: assessing the contribution of the built environment. Prev Med 47:304-8.

Harcourt B.E. (2001) Illusion of Order: The False Promise of Broken Windows Policing Harvard University Press, Boston.

Hill T.D., Ross C.E., Angel R.J. (2005) Neighborhood disorder, psychophysiological distress, and health. J Health Soc Behav 46:170-86.

Jacobs J. (1961) The Death and Life of Great American Cities Vintage Books.

Kelling G.L., Coles C.M. (1998) Fixing Broken Windows: Restoring Order And Reducing Crime In Our Communities Free Press.

Kelly C.M., Schootman M., Baker E.A.*, et al.* (2007) The association of sidewalk walkability and physical disorder with area-level race and poverty. J Epidemiol Community Health 61:978-83.

Kling J.R., Liebman J.B., Katz L.F. (2007) Experimental analysis of neighborhood effects. Econometrica 75:83-119.

Landis JR, Koch GG. (1977) The measurement of observer agreement for categorical data. Biometrics 33:159-174.

McEwen B.S. (1998) Stress, adaptation, and disease - Allostasis and allostatic load, in: S. M. McCann, et al. (Eds.). pp. 33-44.

Miles R. (2006) Neighborhood disorder and smoking: findings of a European urban survey. Soc Sci Med 63:2464-75.

Mohl R.A. (2002) The Interstates and the Cities: highways, housing, and the freeway revolt. Research Report: Poverty and Race Research Action Council.

Molnar B.E., Gortmaker S.L., Bull F.C., Buka S.L. (2004) Unsafe to play? Neighborhood disorder and lack of safety predict reduced physical activity among urban children and adolescents. Am J Health Promot 18:378-86.

Ortiz V., Telles E.E. (2008) Generations of Exclusion: Mexican Americans, Assimilation, and Race Russel Sage Foundation, New York.

Perkins D.D., Taylor R.B. (1996) Ecological assessments of community disorder: Their relationship to fear of crime and theoretical implications. American Journal of Community Psychology 24:63-107.

Perkins D.D., Meeks J.W., Taylor R.B. (1992) The Physical-Environment of Street Blocks and Resident Perceptions of Crime and Disorder - Implications for Theory and Measurement. Journal of Environmental Psychology 12:21-34.

Peterson C.E., Pebley A.R., Sastry N. (2007) The Los Angeles Neighborhood Services and Characteristics Database: Codebook. RAND Working Paper WR-240/12-LAFANS.

Raudenbush S.W., Sampson R.J. (1999) Ecometrics: Toward a science of assessing ecological settings, with application to the systematic social observation of neighborhoods, Sociological Methodology 1999, Vol 29. pp. 1-41.
Ross C.E., Jang S.J. (2000) Neighborhood Disorder, Fear, and Mistrust: The Buffering Role of Social Ties with Neighbors. American Journal of Community Psychology 28:401-420.
Ross C.E., Mirowsky J. (2001) Neighborhood disadvantage, disorder, and health. J Health Soc Behav 42:258-76.
Sampson R.J., Raudenbush S. (1999) Systematic Social Observation of Public Spaces: A new look at disorder in urban neighborhoods. American Journal of Sociology 105:603-651.
Sampson R.J., Raudenbush S.W., Earls F. (1997) Neighborhoods and Violent Crime: A Multilevel Study of Collective Efficacy. Science 277:918-924.
Sastry N., Ghosh-Dastidar B., Adams J., Pebley A.R. (2006) The design of a multilevel survey of children, families, and communities: The Los Angeles Family and Neighborhood Survey. Social Science Research 35:1000-1024.
Skogan W.G. (1990) Disorder and Decline University of California Press, Berkley and Los Angeles, CA.
Social Science Data Analysis Network (SSDAN). (No date) Percent Foreign-Born Ranking, 2000: Metros ranked by percent of population foreign-born. At: http://www.censusscope.org/us/metro_rank_mig_foreign_born.html. Accessed on April 2, 2010.
Taylor R.B., Shumaker S.A., Gottfredson S.D. (1985) Neighborhood-Level Links between Physical Features and Local Sentiments - Deterioration, Fear of Crime, and Confidence. Journal of Architectural and Planning Research 2:261-275.
Wachter S.M., Gillen K.C. (2006) Public Investment Strategies: How They Matter for Neighborhoods in Philadelphia. At: http://www.upenn.edu/penniur/pdf/Public%20Investment%20Strategies.pdf. Accessed on October 23, 2009.
Wei E., Hipwell A., Pardini D., *et al.* (2005) Block observations of neighbourhood physical disorder are associated with neighbourhood crime, firearm injuries and deaths, and teen births. J Epidemiol Community Health 59:904-8.
Wilson J.Q., Kelling G.L. (1982) The Police and Neighborhood Safety: Broken Windows. Atlantic Monthly.
Wilson W.J. (1987) The Truly Disadvantaged: the inner city, the underclass, and public policy University of Chicago Press, Chicago.

Table 1. Description of the L.A.FANS Neighborhood Observation Study Design

Variable	Mean	Std. Dev.	Min	Max
Block faces per block	4.9	3.1	2	35
Blocks per tract	6.5	3.0	2	14
Observations per block face	2.9	0.6	1	7
Observations per block	14.1	8.9	4	92
Observations by tract	91.8	40.4	24	196
Observations per observer	170.5	164.3	4	629

Variable	Total Count
Observations (observer by block face)	5,966
Block faces	2,071
Blocks	422
Tracts	65

Observations per block face	Count	Percent
1	41	2
2	383	18
3	1,464	71
4	158	8
5 – 7	25	1
Total	2,071	100

Table 2. Situational Characteristics for L.A.FANS Neighborhood Observations

Variable	Percent or mean (std. dev.)
Day of visit	
Monday	10%
Tuesday	14%
Wednesday	13%
Thursday	17%
Friday	17%
Saturday	16%
Sunday	14%
Visit time	
Morning	16%
Midday	35%
Afternoon	38%
Evening	11%
Season	
Spring	2%
Summer	7%
Winter	92%
Previous knowledge of block	
None	73%
As part of L.A.FANS study	20%
Outside of study	8%
Minutes spent observing	60.3 (48.8)

Table 3. Block-Face Summary Statistics for L.A.FANS Neighborhood Observations Items

Num	Variable	N	Type[a]	Min	Max	Mean	SD
1	Number of traffic lanes	5937	C	0	9	2.44	1.06
2	Traffic flow	5928	O	1	5	2.10	1.21
4	Street surface	5927	O	1	5	3.22	0.81
5	Sidewalk surface	5939	O	1	6	3.55	1.29
6	Parking restrictions	5896	D	0	1	0.05	0.22
7	Public transportation	5906	D	0	1	0.10	0.30
9	Trees	5924	O	1	4	2.58	1.10
		5924	D	0	1	0.79	0.41
10	Abandoned cars	5941	O	1	4	1.13	0.42
		5941	D	0	1	0.10	0.30
11	Trash or junk	5935	O	1	4	1.84	0.95
		5935	D	0	1	0.53	0.50
12	Garbage, litter, glass	5938	O	1	4	2.22	1.00
		5938	D	0	1	0.73	0.44
13	Drug paraphernalia	5936	O	1	4	1.04	0.22
		5936	D	0	1	0.03	0.18
14	Beer or liquor bottles	5938	O	1	4	1.29	0.62
		5938	D	0	1	0.21	0.41
15	Cigarette butts	5937	O	1	4	1.94	0.97
		5937	D	0	1	0.60	0.49
16	Graffiti	5935	O	1	4	1.88	1.00
		5935	D	0	1	0.53	0.50
17	Painted-over graffiti	5938	O	1	4	1.52	0.79
		5938	D	0	1	0.37	0.48
18	Strong odors	5913	D	0	1	0.06	0.24
19	Land use: residential	5930	D	0	1	0.77	0.42
	Land use: residential/ commercial	5930	D	0	1	0.08	0.28
	Land use: other	5930	D	0	1	0.03	0.16
20	Housing type: stand alone houses	5918	D	0	1	0.79	0.41
	Housing type: duplexes	5918	D	0	1	0.18	0.39
	Housing type: multiple occupancy	5918	D	0	1	0.17	0.37
	Housing type: low-rise	5918	D	0	1	0.23	0.42
	Housing type: mid-rise	5918	D	0	1	0.02	0.14
21	Overall condition of residential buildings	5339	O	1	5	2.49	0.92
22	Residential: burned, boarded, abandoned	5346	O	1	5	1.11	0.37
		5346	D	0	1	0.09	0.29
23	Vacant lots	5351	O	1	5	1.21	0.56
		5351	D	0	1	0.16	0.36
24	Damaged exteriors	5348	O	1	5	2.16	1.06
		5348	D	0	1	0.68	0.47
25	Well-tended yards	5346	O	1	5	2.66	1.28
		5346	D	0	1	0.78	0.42
26	Residential: window bars	5339	O	1	5	2.46	1.35
		5339	D	0	1	0.64	0.48

27	Residential: private security	5343	O	1	5	1.86	1.04
		5343	D	0	1	0.51	0.50
28	Residential: dogs	5350	O	1	5	1.42	0.70
		5350	D	0	1	0.32	0.47
29	Residential: security fences	5348	O	1	5	2.28	1.32
		5348	D	0	1	0.59	0.49
30	Neighborhood watch	5350	D	0	1	0.17	0.38
31	Residential: for sale/rent	5352	O	1	5	1.36	0.63
		5352	D	0	1	0.29	0.45
32	Old, beat-up cars	5349	O	1	4	1.37	0.65
		5349	D	0	1	0.28	0.45
33	Commercial: overall condition	1311	O	1	5	3.32	0.86
34	Commercial: abandon, burned, boarded	1316	O	1	5	1.23	0.65
		1316	D	0	1	0.15	0.36
35	Commercial: window bars	1314	O	1	5	2.07	1.45
		1314	D	0	1	0.44	0.50
36	Commercial: security fences	1314	O	1	5	2.54	1.57
		1314	D	0	1	0.60	0.49
37	Commercial: for sale/rent	1312	O	1	5	1.18	0.52
		1312	D	0	1	0.13	0.34
40	Public telephone	5922	D	0	1	0.12	0.32

Note

[a] Type of variable: C = continuous; O = ordinal; D = dichotomous.

Table 4. Block-Face and Tract Intra-Class Correlation Coefficients across Multiple Observers for L.A.FANS Neighborhood Observations Items

			Model 1	Model 2	Model 3	Model 4	Model 5	
	Situational variables[c]		No	Yes	No	Yes	Yes	Yes
#	**Variable**	**Type[a]**	**Block-face ICC**	**Block-face ICC[b]**	**Tract ICC[b]**	**Tract ICC[b]**	**Tract ICC[b]**	**Block-face ICC[b]**
1	Num of traffic lanes	C	0.86	0.86	0.11	0.10	0.08	0.78
2	Traffic flow	O	0.74	0.74	0.16	0.15	0.17	0.59
4	Street surface	O	0.47	0.46	0.27	0.30	0.30	0.19
5	Sidewalk surface	O	0.70	0.67	0.54	0.53	0.53	0.16
6	Parking restrictions	D	0.72	0.69	0.57	0.57	0.50	0.28
7	Public transportation	D	0.87	0.87	0.17	0.18	0.04	0.87
9	Trees	O	0.52	0.53	0.21	0.22	0.23	0.32
		D	0.60	0.61	0.23	0.25	0.26	0.39
10	Abandoned cars	O	0.52	0.50	0.47	0.45	0.42	0.18
		D	0.49	0.48	0.47	0.45	0.40	0.14
11	Trash or junk	O	0.43	0.42	0.28	0.26	0.27	0.17
		D	0.46	0.43	0.29	0.29	0.30	0.16
12	Garbage, litter, glass	O	0.43	0.42	0.31	0.30	0.40	0.16
		D	0.54	0.50	0.42	0.41	0.39	0.12
13	Drug paraphernalia	O	0.43	0.42	0.42	0.47	0.43	0.12
		D	0.44	0.43	0.42	0.48	0.43	0.13
14	Beer or liquor bottles	O	0.34	0.33	0.35	0.34	0.34	0.08
		D	0.36	0.34	0.35	0.35	0.35	0.08
15	Cigarette butts	O	0.38	0.34	0.32	0.29	0.30	0.09
		D	0.47	0.40	0.39	0.38	0.38	0.11
16	Graffiti	O	0.66	0.61	0.55	0.53	0.53	0.11
		D	0.79	0.73	0.63	0.62	0.61	0.12
17	Painted-over graffiti	O	0.49	0.43	0.48	0.49	0.49	0.06
		D	0.54	0.46	0.50	0.52	0.53	0.06
18	Strong odors	D	0.32	0.29	0.21	0.20	0.19	0.14
19	Land use: residential	D	0.86	0.87	0.18	0.18	0.22	0.63
	Land use: other	D	0.68	0.70	0.30	0.31	0.20	0.50
20	Housing type							
	Stand-alone houses	D	0.95	0.93	0.26	0.26	0.25	0.62
	Duplexes	D	0.53	0.48	0.50	0.45	0.43	0.15
	Multiple occupancy	D	0.58	0.54	0.51	0.50	0.47	0.18
	Low-rise	D	0.82	0.80	0.47	0.45	0.44	0.38
	Mid-rise	D	0.77	0.77	0.63	0.63	0.54	0.24
21	Res: condition	O	0.64	0.56	0.54	0.55	0.55	0.08
22	Res: abandoned, burned, boarded	O	0.68	0.66	0.42	0.42	0.36	0.34
		D	0.68	0.66	0.42	0.42	0.36	0.34
23	Vacant lots	O	0.81	0.81	0.41	0.41	0.38	0.44
		D	0.81	0.82	0.41	0.41	0.38	0.44
24	Damaged exteriors	O	0.48	0.46	0.37	0.37	0.38	0.12
		D	0.57	0.53	0.40	0.40	0.40	0.16
25	Well-tended yards	O	0.59	0.56	0.42	0.42	0.44	0.15

		D	0.62	0.56	0.50	0.51	0.50	0.15
26	Res: window bars	O	0.68	0.62	0.56	0.56	0.56	0.07
		D	0.72	0.64	0.60	0.60	0.60	0.10
27	Res: private security	O	0.61	0.57	0.35	0.35	0.35	0.25
		D	0.58	0.55	0.29	0.29	0.30	0.30
28	Res: Dogs	O	0.55	0.54	0.22	0.22	0.22	0.36
		D	0.57	0.57	0.22	0.22	0.21	0.37
29	Res: security fences	O	0.38	0.37	0.23	0.26	0.26	0.12
		D	0.37	0.37	0.25	0.30	0.31	0.12
30	Neighborhood watch	D	0.65	0.66	0.17	0.17	0.14	0.50
31	Res: for sale/rent	O	0.67	0.67	0.16	0.16	0.16	0.52
		D	0.73	0.72	0.15	0.15	0.16	0.57
32	Old, beat-up cars	O	0.49	0.45	0.33	0.33	0.33	0.18
		D	0.45	0.41	0.32	0.32	0.33	0.15
33	Comm: condition	O	0.47	0.49	0.33	0.33	0.33	0.16
34	Comm: abandoned, burned, boarded	O	0.59	0.59	0.25	0.24	0.16	0.44
		D	0.58	0.59	0.24	0.23	0.16	0.44
35	Comm: window bars	O	0.39	0.39	0.28	0.28	0.28	0.15
		D	0.46	0.45	0.31	0.30	0.30	0.19
36	Comm: security fences	O	0.51	0.52	0.21	0.23	0.23	0.30
		D	0.47	0.51	0.22	0.25	0.25	0.28
37	Comm: for sale/rent	O	0.66	0.68	0.19	0.19	0.10	0.60
		D	0.67	0.70	0.20	0.20	0.12	0.59
40	Public telephone	D	0.82	0.81	0.38	0.38	0.32	0.51
	Situational variables[c]		No	Yes	No	Yes	Yes	Yes

Notes

[a] Type of variable: C = continuous; O = ordinal; D = dichotomous.

[b] ICC = intra-class correlation coefficients, which are based on models with random effects. Linear regression models are used for continuous variables, ordered logit models for ordered outcomes with more than two response categories, and logit models for dichotomous outcomes.

[c] Situational variables include duration of visit, day of visit, time of visit, season, and previous knowledge of block.

[d] Abbreviations: "res" = residential and "comm" = commercial.

Table 5. Multilevel Item Response Models of Neighborhood Physical Characteristics Based on the L.A.FANS Neighborhood Observations Items

Variable	Model 1 Seven-item scale		Model 2 Twenty-item scale	
Observation items				
10 Abandoned cars	–3.17***	(0.21)	–2.82***	(0.17)
11 Trash or junk	–	–	0.27	(0.17)
12 Garbage, litter, glass	1.72***	(0.21)	1.57***	(0.17)
13 Drug paraphernalia	–4.51***	(0.22)	–4.10***	(0.18)
14 Beer or liquor bottles	–1.98***	(0.21)	–1.72***	(0.17)
15 Cigarette butts	0.70***	(0.21)	0.66***	(0.17)
16 Graffiti	0.23	(0.21)	0.25	(0.17)
17 Painted-over graffiti	–0.86***	(0.21)	–0.71***	(0.17)
18 Strong odors	–	–	–3.43***	(0.17)
20/3 Stand-alone houses	–	–	–1.97***	(0.17)
20/4 Duplexes	–	–	–2.12***	(0.17)
21 Residential: condition	–	–	2.62***	(0.17)
22 Res: abandoned, burned, boarded	–	–	–2.91***	(0.17)
24 Damaged exteriors	–	–	1.20***	(0.17)
25 Well-tended yards	–	–	1.94***	(0.17)
26 Residential: window bars	–	–	0.96***	(0.16)
28 Residential: Dogs	–	–	–0.99***	(0.17)
32 Old, beat-up cars	–	–	–1.23***	(0.17)
33 Commercial: condition	–	–	3.58***	(0.25)
35 Commercial: window bars	–	–	–1.00***	(0.18)
Observation duration (mins.)	–0.002**	(0.001)	–0.001**	(0.0005)
Day of week				
Monday	–0.8	(0.07)	–0.05	(0.04)
Tuesday	–0.11*	(0.07)	–0.09**	(0.04)
Wednesday	0.30***	(0.06)	0.22***	(0.04)
Thursday[a]	.	.	.	.
Friday	0.12*	(0.06)	0.01	(0.04)
Saturday	0.17***	(0.06)	0.11***	(0.04)
Sunday	0.26***	(0.06)	0.10**	(0.04)
Time of day				
Morning	–0.005	(0.05)	0.02	(0.03)
Midday	0.13***	(0.04)	–0.02	(0.03)
Afternoon[a]	.	.	.	.
Evening	0.04	(0.06)	0.02	(0.03)
Season				
Spring[a]	.	.	.	.
Summer	0.75***	(0.14)	0.47***	(0.09)
Winter	0.72***	(0.08)	0.26***	(0.05)
Interviewer knows neighborhood				
No[a]	.	.	.	.
Through L.A.FANS study	0.33***	(0.05)	0.23***	(0.03)
Other experience	–0.002	(0.07)	0.04	(0.04)

Variance of random effects				
Block-face	0.89***	(0.05)	0.57***	(0.03)
Tract	2.37***	(0.43)	1.63***	(0.29)
Model Chi-squared	8,141***	(df=21)	23,171.51	(df=34)
Observations				
Items by observer by block face	41,762		105,289	
Observations by block face	5,966		5,966	
Block faces	2,071		2,066	
Tracts	65		65	

Note: [a] Reference category

Table 6. Summary Measures for Tract-Level Disorder Scales Based on L.A.FANS Neighborhood Observations

Measure	Seven-item scale	Twenty-item scale
ICC		
Block-face	0.14	0.10
Tract	0.36	0.30
Summary statistics		
Min	–2.71	–2.96
Max	2.35	1.74
Mean	0.00	0.00
SD	1.54	1.28

Table 7. Linear Regression Models of L.A.FANS Tract Disorder Score on Neighborhood Structural Characteristics

Variable	**Model 1 Seven-item scale**		**Model 2 Twenty-item scale**	
Constant	–0.33	(0.22)	–0.17	(0.19)
Concentrated disadvantage	0.61***	(0.13)	0.46***	(0.11)
Concentrated affluence	–0.42*	(0.23)	–0.76***	(0.19)
Immigrant concentration	0.21	(0.20)	–0.23	(0.17)
Residential stability	–0.20*	(0.12)	–0.20**	(0.10)
Ethnic diversity	–0.79*	(0.42)	–0.80**	(0.33)
Model F-statistic	86.65***	(df = 5, 59)	83.75***	(df=5, 59)
Adjusted R-squared	0.87		0.87	
Observations	65		65	