Systematic Social Observation of Public Spaces: A New Look at Disorder in Urban Neighborhoods

Citation	Sampson, Robert J. and Stephen W. Raudenbush. 1999. Systematic social observation of public spaces: A new look at disorder in urban neighborhoods. American Journal of Sociology 105, no. 3: 603-651.
Published Version	http://dx.doi.org/10.1086/210356
Citable link	http://nrs.harvard.edu/urn-3:HUL.InstRepos:3226951
Terms of Use	This article was downloaded from Harvard University's DASH repository, and is made available under the terms and conditions applicable to Other Posted Material, as set forth at http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA

Systematic Social Observation of Public Spaces: A New Look at Disorder in Urban Neighborhoods[1]

Robert J. Sampson
University of Chicago and American Bar Foundation

Stephen W. Raudenbush
University of Michigan

This article assesses the sources and consequences of public disorder. Based on the videotaping and systematic rating of more than 23,000 street segments in Chicago, highly reliable scales of social and physical disorder for 196 neighborhoods are constructed. Census data, police records, and an independent survey of more than 3,500 residents are then integrated to test a theory of collective efficacy and structural constraints. Defined as cohesion among residents combined with shared expectations for the social control of public space, collective efficacy explains lower rates of crime and observed disorder after controlling neighborhood structural characteristics. Collective efficacy is also linked to lower rates of violent crime after accounting for disorder and the reciprocal effects of violence. Contrary to the "broken windows" theory, however, the relationship between public disorder and crime is spurious except perhaps for robbery.

> The answer to the question of how city life was to be possible, then, is this. *City life was made possible by an "ordering" of the urban populace in terms of appearance and spatial location such that those within the city could know a great deal about one another by simply looking.*
>
> (Lyn Lofland, *A World of Strangers: Order and Action in Urban Public Space,* 1973, p. 22; emphasis in original)

Visual signs of social and physical disorder in public spaces reflect powerfully on our inferences about urban communities. By social disorder, we refer to behavior usually involving strangers and considered threatening,

[1] We thank Tony Earls, Albert J. Reiss Jr., Steve Buka, Jeffrey Morenoff, Richard Congdon, and Matheos Yosef for their help in this project, and the NORC team led by Woody Carter, Cindy Veldman, Jody Dougherty, and Ron Boyd for heroic efforts in data collection. John Laub and the *AJS* reviewers provided helpful comments on an earlier draft. The long-standing interest of Albert J. Reiss, Jr., in systematic social

0002-9602/2000/10503-0001$02.50

such as verbal harassment on the street, open solicitation for prostitution, public intoxication, and rowdy groups of young males in public. By physical disorder, we refer to the deterioration of urban landscapes, for example, graffiti on buildings, abandoned cars, broken windows, and garbage in the streets. Visible evidence of disorder, or what Albert Hunter (1985) calls "incivilities," have long been noted as central to a neighborhood's public presentation (Goffman 1963). Jane Jacobs's classic observation of urban life in the 1950s even then evoked a concern with the threats of disorder to neighborhood civility (1961, pp. 29–54), especially the negotiation of public encounters in the "world of strangers" (Lofland 1973).[2]

The streets, parks, and sidewalks still belong to no one and therefore to everyone. Disorder continues to be of theoretical interest precisely because of its visual salience and symbolism regarding the use of such spaces. Even if we wish it were not so, disorder triggers attributions and predictions in the minds of insiders and outsiders alike. It changes the calculus of prospective home buyers, real estate agents, insurance agents, and investors and shapes the perceptions of residents who might be considering moving. Evidence of disorder also gives a running account of the effectiveness of residents seeking neighborhood improvement, and that record may encourage or discourage future activism. Physical and social disorder in public spaces are thus fundamental to a general understanding of urban neighborhoods.

Neighborhood disorder has more specific bearing on the study of crime as well. Research has established connections between disorder and both fear of crime and crime rates (Skogan 1990; Kelling and Coles 1996). In fact, a reigning theory posits that minor disorder is a direct cause of serious crime. Originators of the "broken windows" thesis, Wilson and Kelling (1982) argued that public incivilities—even if relatively minor as in the case of broken windows, drinking in the street, and graffiti—attract predatory crime because potential offenders assume from them that residents are indifferent to what goes on in their neighborhood. The metaphor of

observation provided the inspiration and practical guidance for our work, and the Center for Advanced Study in the Behavioral Sciences provided a most hospitable intellectual environment for its fruition. Data collection was funded in part by the John D. and Catherine T. MacArthur Foundation and the National Institute of Justice. Direct correspondence to Robert J. Sampson, Department of Sociology, University of Chicago, 1126 East Fifty-ninth Street, Chicago, Illinois 60637. E-mail: rjsam@src.uchicago.edu

[2] Goffman (1963) goes back yet further, to the obligation in medieval times to keep one's pigs out of the streets. In this case, the norms regulating public order refer not just to face-to-face interaction among strangers or acquaintances, but the visual ordering of the physical landscape (1963, p. 9). For example, Goffman writes of expectations regarding the maintenance of sidewalks and keeping the streets free of refuse. Hence, disorder in public places may be conceived in physical as well as social terms.

broken windows is apt, insofar as the theory asserts that physical signs of disorder serve as a signal of the unwillingness of residents to confront strangers, intervene in a crime, or call the police (Greenberg and Rohe 1986; Skogan 1990, p. 75). Proponents thus assume that both physical disorder and social disorder provide important environmental cues that entice potential predators. The "broken windows" thesis has gained ascendancy in criminology and has greatly influenced public policy, leading to police crackdowns in numerous cities on the manifestations of social and physical disorder. New York City is the best-known example of aggressive police tactics to control public incivilities (Kelling and Coles 1996, pp. 108–56; Kelling 1998).

Taking seriously the idea that visual cues matter, this article applies the method of systematic social observation (SSO) to the study of social and physical disorder in urban neighborhoods. We depart from prior research in three ways. First, we describe novel systematic procedures for collecting observational assessments of public spaces using videotaping procedures that produce a permanent visual record amenable to later coding and reinterpretation based on emergent insights.[3] Second, we formulate a hierarchical item-response model that identifies sources of error in aggregating across observed disorder items within block faces and in aggregating across block faces within some 200 census tracts. The method yields high tract-level reliabilities for assessing both social disorder and physical disorder.

The third and major goal of the article is to assess the sources and consequences of neighborhood disorder. We do so by testing the association of systematically observed disorder with independent measures of officially recorded and survey-reported crime, census-based sociodemographic composition, and a survey-based measure that taps the collective efficacy of residents in achieving informal social control. A theory combining structural constraints with local collective efficacy is presented as an alternative to the "broken windows" interpretation of the disorder-crime link. We also assess broader implications for reinvigorating the study of urban communities based on systematic observation and video-based approaches.

SYSTEMATIC SOCIAL OBSERVATION OF DISORDER

In the spirit of the early Chicago school of urban sociology, we believe that direct observation is fundamental to the advancement of knowledge

[3] Building on Lofland (1973, p. 19), we define public spaces as those areas of a neighborhood to which persons have legal access and can visually observe—its streets and sidewalks, its parks, its places of public accommodation, its public buildings, and the public sectors of its private buildings.

(Park and Burgess 1921; see also Whyte 1988). As Andrew Abbott (1997) notes, one of the hallmarks of the Chicago school was its concern with observing public places—not just abstract variables, but the sights, sounds, and feel of the streets. Attempting to systematize such approaches, more than 25 years ago Albert J. Reiss, Jr. (1971) advocated systematic social observation as a key measurement strategy for natural social phenomena. By systematic, Reiss meant that observation and recording are done according to explicit rules that permit replication; he also argued that the means of observation, whether a person or technology, must be independent of that which is observed. By natural social phenomena, Reiss (1971, p. 4) meant "events and their consequences, including properties of organization, can be observed more or less as they occur." As his main example, Reiss described systematic observations of police-citizen encounters. He also noted the general import of the SSO method for assessing physical conditions and social interactions within neighborhood settings that survey respondents may be incapable of describing accurately.

Despite the potential yield of direct observation, the majority of research studies linking signs of disorder with fear of crime and criminal victimization have been based on residents' subjective perceptions drawn from survey responses.[4] The typical strategy in survey research has been to ask residents how much of a problem they perceive disorder to be; the standard finding is that perceptions of disorder predict fear of crime (Skogan 1990; Perkins and Taylor 1996; Taylor 1997, 1999). The dearth of independent assessments of neighborhood disorder poses a special problem for interpreting this linkage. As Taylor (1999) has argued, the high correlation between fear and disorder may arise in part from shared survey-method variance. More fundamentally, however, the perception of disorder seems also to reflect a psychological construct—perhaps fear itself (Garofalo and Laub 1978; Rountree and Land 1996). Residents fearful of crime report more disorder than do residents who experience less fear, even though both sets of observers are reporting on the *same* neighborhood (Taylor 1999; Perkins, Meeks, and Taylor 1992). In this scenario, the fear (or vulnerability) of residents might be said to induce perceptions of disorder. Even the disorder-crime link is problematic, since victimization experiences are usually measured in the same surveys used to assess (perceived) disorder.

One of the primary obstacles to bringing independent and systematic social observation to bear on this conundrum has been methodological uncertainty, not just on how to properly conduct such observations, but on how to properly assess their measurement properties at the neighbor-

[4] Useful reviews of the empirical literature are found in Perkins and Taylor (1996), Taylor (1997, 1999), Skogan (1990), and Skogan and Maxfield (1981).

hood level (Raudenbush and Sampson 1999*b*). Another concern has been cost, even though direct observations are potentially less expensive than household surveys, with listing, screening, broken appointments, and response rates eliminated. Yet another obstacle has been conceptual, stemming from underappreciation of the yield of systematic observation for one of the fundamental cleavages in sociological criminology—the reliable and valid measurement of crime and deviance. Perhaps most important, however, has been the psychological reductionism that flows from the dominant theoretical and empirical focus on individuals.

An exception to the lack of independent observations of disorder at the level of ecological units rather than persons is found in research by Taylor and colleagues (Taylor, Shumaker, and Gottfredson 1985; Taylor, Gottfredson, and Brower 1984; Covington and Taylor 1991; Perkins et al. 1992; Perkins and Taylor 1996; see also Mazerolle, Kadleck, and Roehl 1998).[5] Using observations conducted by teams of trained raters walking the streets, Taylor et al. (1985) assessed 20% of all occupied face blocks in 66 Baltimore neighborhoods. A face block is the block segment on one side of a street. They identified two physical dimensions that stood out empirically: physical incivilities and nonresidential land use. These two dimensions were reliable in terms of individual-level psychometrics (e.g., Cronbach's alpha; interrater reliability) and were related as expected to perceived disorder and fear of crime derived from neighborhood surveys. More recently, Perkins et al. (1992) examined on-site assessments of block-level physical incivilities in Baltimore. Controlling for social factors, physical incivilities predicted perceptions of crime-related problems. Yet, using similar procedures in a different city, Perkins et al. (1993) report that residents' perceptions and an independent rating of physical disorder were not significantly correlated. Observed environmental items correlated more strongly with multiple indicators of subsequent block crime than did residents' perceptions of the environment. Interestingly, residents' perceptions of physical disorder correlated positively with fear, but not after controls were introduced for income, stability, and racial composition.

Overall, then, the research record is mixed and curiously imbalanced. Although specified as an ecological construct, neighborhood disorder has been investigated mainly using individual perceptions and individual-level research designs. The number of studies employing observational ratings across multiple ecological contexts is small, and the correlation of observed disorder with subjective perceptions varies by level of aggrega-

[5] An early version of systematic observation based on single interviewer ratings in a neighborhood survey was used in Taub, Taylor, and Dunham (1984) and Skogan (1990).

tion, type of measure, and study site. We therefore approach the study of disorder from an integrated observational, survey, and record-based approach at the neighborhood level. Just as important, we offer a theoretical framework on the sources and consequences of public disorder that challenges the prevailing view on broken windows and crime.

Rethinking Disorder

Rather than conceive of disorder as a direct cause of crime, we view many elements of disorder as part and parcel of crime itself. Consider the typical items used to define social disorder, such as solicitation for prostitution, loitering, and public use of alcohol or drugs. Consider also physical "incivilities," such as graffiti, smashed windows, and drug vials in the streets. All of these are evidence either of crimes (whether concurrent with the observation, as in drug use, or physical evidence of recent acts, as in drug vials on the sidewalk) or ordinance violations. Although ordinance violations like drinking in public and many "soft crimes" like graffiti may not be judged as particularly serious, this is an evaluation or classification issue and not a statement on etiology. As Gottfredson and Hirschi (1990, pp. 42–43) have argued, criminologists often mistake differences in crime seriousness (classification) for differences in causal mechanisms. But as the long history of developmental research on juvenile delinquency has instructively shown, minor offenses usually do the best job of discriminating individual differences in later serious crime. In fact, early smoking and truancy are among the most reliably measured predictors of the propensity to serious adolescent delinquency (Farrington 1979), presumably because these acts all have the same antecedents (Gottfredson and Hirschi 1990). Relatedly, Hagan and McCarthy's (1997) recent ethnographic and quantitative study of two Canadian cities demonstrates the close connection of predatory youth crime to street life and settings of public disorder (e.g., prostitution, vagrancy, drug selling).

Applying the logic of Gottfredson and Hirschi (1990) to the present neighborhood-level case, a reasonable hypothesis is that public disorder and predatory crimes are manifestations of the same explanatory process, albeit at different ends of a "seriousness" continuum. Even those elements of disorder not obviously criminal in nature (e.g., garbage, vacant housing) are either violations of an ordinance (as in littering, slumlord abandonment) or may be conceptualized as sharing a similar causal structure and thus predicted by similar mechanisms (Hunter 1985). Concretely, for example, it does not seem to us persuasive to argue that graffiti causes robbery. Lack of social control might cause both graffiti and robbery; if so, one should measure the specified causal mechanism rather than (tautologically) inferring lack of order from graffiti and then using it to explain robbery. What makes this conceptual move significant, in our view, is that it provides the opportunity to *observe* and hence systematically measure important manifestations of crime-related processes. Muggings, assaults, and rapes might be impossible to reliably observe, but vandalism, prostitution, gang congregation, and evidence of drug use can, in principle, be observed by all, whether residents, business people, visitors, possible investors, local activists, or potential offenders. Sociologists of crime have debated for at least 30 years the relative merits of survey-reported crime (whether offending or victimization) relative to official police records (Short and Nye 1957). By recasting disorder in the theoretical terms of crime, an observational window is opened on a new alternative for testing neighborhood-level theory.

Of course, not all environmental observations tap disorder. Researchers have profitably examined land use (e.g., mixed residential-business), the presence of bars, street layout, traffic patterns, and housing structure, all of which are conceptually distinct from crime (e.g., Perkins and Taylor 1996; Taylor 1999; Perkins et al. 1992). Such environmental features are legitimate targets of observation, and we too incorporate them in our methodology. Moreover, our argument is not that social and physical disorder are unimportant for explaining neighborhood dynamics. To the contrary, our framework suggests that while both crime and disorder reflect common origins, crime may be less relevant for understanding processes such as population abandonment and the perceived incivility of urban life because it is largely unobserved (see also Jacobs 1961; Skogan 1990). Corresponding to the "text" from which all key actors in a neighborhood read, we propose that disorder is the more visually proximate or immediate neighborhood cue of theoretical interest, even if it is not a direct cause of further crime.

STRUCTURAL CONSTRAINTS AND THE AGENCY OF SOCIAL CONTROL

The theoretical framework that guides our assessment of disorder stems from a balancing of structural constraints with recognition of purposive social action. In the study of crime and disorder, attention has focused primarily on structural dimensions of an economic nature over which residents are thought to have little control—especially concentrated poverty and its associated lack of social resources. The ecological concentration of disadvantage means that the poorest neighborhoods tend to have not only the lowest incomes but also higher rates of unemployment, financial dependence, and institutional disinvestment (Wilson 1987; Land, McCall, and Cohen 1990; Hagan and Peterson 1995). Economic deprivation is relevant in that repairing buildings and cleaning up residential and commercial areas requires money. Because areas of concentrated disadvantage find it difficult to support viable commercial enterprise, many stores and apartments will be vacant, giving little incentive for investors to repair their properties. The density of children in single-parent families, which is ecologically concentrated with poverty and resource dependence in U.S. cities (Land et al. 1990), adds another layer of difficulty to the always challenging task of supervising children and adolescent peer groups.

Structural constraints are not just economic in nature. Systemic theories of urban communities (Kasarda and Janowitz 1974) have long pointed to the importance of residential stability as a major feature of urban social organization. High levels of home ownership and low transience work together to instill in residents a "stake in conformity," in this case to neighborhood well-being. The formation of social networks that undergird local ties and attachment to place is also linked to residential stability (Sampson 1988; Taylor 1997). By stability we do not mean lack of change but rather the social reproduction of neighborhood residential structure, typically when population gains offset losses and home values appreciate.

Against the backdrop of resources and stability, a number of other structural constraints impinge on the ability of neighborhoods to counteract public incivilities, including the sheer density of population, nonresidential land use, public transportation nodes, and large flows of population that overwhelm local services. The "routine activities" perspective (Cohen and Felson 1979) builds on the insight that predatory crime involves the intersection in time and space of motivated offenders, suitable targets, and the absence of capable guardians. Because illegal activities feed on the spatial and temporal structure of routine legal activities (e.g., transportation, work, and shopping), the differential land use of cities is a key to comprehending neighborhood crime, and, by implication, disorder patterns. The effects of concentrated resource disadvantage and residential instability on disorder should thus be considered in concert with structural characteristics such as density, street activity, and commercial land use.

Structural constraints notwithstanding, one might view human agency as central to the explanation of disorder. In this view, it is not only the material circumstances or ecological structures that residents face, but the challenge to organize themselves to achieve shared public ends. We adopt the formulation of Janowitz (1975, pp. 82, 87) and refer to social control as the capacity of a social unit to regulate itself according to desired principles—to realize collective, as opposed to forced, goals. Hence, social control should not be equated with repression or forced conformity. Similar to Bursik (1988), our strategy also highlights variations in social control across ecological units rather than elevating solidarity or affective identity to the major definitional criteria of neighborhood (see also Tilly 1973, p.

212). When formulated in this way, the dimensions of social control are analytically separable not only from possible structural antecedents (e.g., poverty, instability) and effects (e.g., disorder, crime) but from the definition and operationalization of the units of analysis.

Building on Janowitz's (1975) conception of social control requires that we explicitly note the assumption of relative consensus; namely, one of the most central of common goals is the desire of community residents to live in safe environments free of predatory crime and disorder. There is no evidence of which we are aware showing public approval of crime or disorder by any population group (Kornhauser 1978; Hearn 1997); if anything, the evidence suggests that residents of low-income, African-American, and high-crime neighborhoods are the *most* insistent on better police protection and demands for reducing violence (Skogan and Hartnett 1998). Although existential weariness in the inner city may lead to a greater tolerance of certain forms of deviance, it is precisely the acceptance of common standards by residents and even gang leaders themselves that underlies efforts to establish social order and safety—however unconventional those efforts may be (Sampson and Jeglum-Bartusch 1998). Indeed, Pattillo's (1998) revealing ethnography of a black middle-class neighborhood on the south side of Chicago found that the incorporation of gang members and drug dealers into the networks of law-abiding kin and neighbors thwarted efforts to rid the neighborhood of its criminal element. Yet in an interesting twist, Pattillo found that the leader of a major black gang was a long-time resident who engaged in multiple acts of social control (e.g., threats, monitoring) to keep the neighborhood free of street crime and signs of disorder (e.g., graffiti, vandalism, prostitution). Pattillo writes that both sides—the residents and the gang leaders—"spurn disorder, actively combat graffiti, and show disdain for activities that may invite negative attention, such as loitering or public fighting" (1998, p. 755). Her findings also highlight the important distinction between "crime" and "criminals"—ironically, even active participants in criminal networks seek to achieve some semblance of order in their neighborhoods of residence. The phenomenon of informal efforts to socially control local crime and disorder has long been reported in white working-class neighborhoods dominated by the mob (e.g., Whyte 1943).[6]

[6] Heeding Whyte (1943), we acknowledge that disorder is in some respects a misleading term if not properly contextualized, for obviously many elements of disorder do not involve disorganization or a chaotic pattern. Prostitution and drug dealing, e.g., may follow quite explicit rules of street organization (Pattillo 1998). By disorder, then, we refer not to disorganization but observable physical and social cues that are commonly perceived to disturb the civil and unencumbered use of public space (see Skogan 1990; Hearn 1997). Moreover, we focus on the ecological *concentration* of *multiple* dimensions of public disorder. A particular exhibit of graffiti might be considered an artistic expression, but the recurrent defacement of public property accompa-

In contrast to externally or formally induced actions (e.g., a police crackdown, housing code enforcement), our agency-oriented perspective on neighborhoods emphasizes the role of informal mechanisms by which residents initiate or achieve social control. Examples of indigenous informal social control relevant to reducing disorder include the willingness of residents to intervene to prevent acts such as truancy, drinking, vandalism, and street-corner disturbances (e.g., harassment, loitering, fighting). Importantly, however, actions of informal control need not involve direct confrontation or exclude the police or other formal channels of recourse. Recent ethnographic research has identified the creative ways in which socially organized communities react to disorder, including the establishment of "phone trees" among residents for calling the police upon observation of disorder; the organization of a group presence in court sentencing hearings for offenders caught defacing local public properties; voluntary "graffiti patrols" that log new incidents of disorder that are then presented to the police; and agitating for voting referendums to delicense bars where drug sales and disorder loom large (Carr 1998). The razing of a vacant "drug house" by housing authorities, if prompted by local complaints, would also fit this pattern. Ultimately, then, our perspective recognizes the articulation among the private (family), parochial (neighborhood), and formal (public) orders (Hunter 1985; Bursik and Grasmick 1993) but stresses the agency of residents in establishing these connections.

In short, we theorize informal social control as a dynamic process differentially activated across neighborhoods. Drawing on Sampson, Raudenbush, and Earls (1997), we propose an analogy between individual efficacy and neighborhood efficacy: both refer to the capacity for achieving an intended effect. At the neighborhood level, the shared willingness of local residents to intervene for the common good depends, in addition, on conditions of cohesion and mutual trust among neighbors. One is unlikely to take action in a neighborhood context where the rules are unclear and people mistrust one another. Personal ties and friendship are not sufficient; the private world of strong kinship ties may actually interfere with public trust and the expectation of collective responsibility for getting things done (Whyte 1943; Jacobs 1961, p. 82; Carr 1998). Attempting to transcend the traditional focus in urban systemic theory on personal ties, we thus define "collective efficacy" as the linkage of cohesion and mutual trust with shared expectations for intervening in support of neighborhood

nied by other physical signs of deterioration (e.g., broken windows, abandoned cars, trash in the streets) signals to most observers a perceived threat to public order. Neighborhood disorder is thus in the end a useful concept, not the least because it conjures up powerful visual images that cut across all population subgroups.

social control (Sampson et al. 1997).[7] Just as individuals vary in their capacity for efficacious action, so too do neighborhoods. And just as individual self-efficacy is situated relative to a particular task rather than global, our notion of collective efficacy here is conceptualized as relative to the task of maintaining order in public spaces.

To be sure, structural constraints and process-oriented mechanisms such as informal social control are not mutually exclusive. A more plausible account would view structural constraints and human agency as interrelated, jointly and reciprocally shaping social action (Sewell 1992). Within the limitations of our data, we therefore expect simultaneous contributions of structural characteristics and collective efficacy to the explanation of observed disorder and crime. Theory elaborated elsewhere (Sampson et al. 1997, p. 919) also suggests that concentrated resource disadvantage and residential instability are major structural conditions that undermine collective efficacy, in turn fostering increased crime and, by implication, public disorder. A theory of collective efficacy thus does not render structural constraints irrelevant; rather, it proposes mediating mechanisms while at the same time insisting on an independent role for agency in all corners of the social structure. We provide a strong test of the theory of collective efficacy by considering as well the simultaneous or reciprocal effect of crime itself on residents' sense of mutual trust and shared expectations for social control (Skogan 1990).

Research Strategy

Sharing an affinity with routine activity theory (Cohen and Felson 1979; Cohen, Kluegel, and Land 1981; Felson 1987; Brantingham and Brantingham 1995), the logic of our analytic approach diverges from a concern with the production of offenders as in the classic social-disorganization tradition of Shaw and McKay (1942; see also Bursik 1988). In the modern urban system, residents traverse the boundaries of multiple neighborhoods during the course of a day (Felson 1987), a problematic scenario for neighborhood theories seeking to explain contextual effects on individual differences in offending. By contrast, we are interested in how neighborhoods fare as units of control or guardianship over their own public spaces (Cohen et al. 1981)—regardless of where offenders may reside. The unit

[7] Note the affinity with Jacobs's (1961, p. 119) focus on the "self-government functions of city streets: to weave webs of public surveillance and thus to protect strangers as well as themselves; to grow networks of small-scale, everyday public life and thus of trust and social control; and to help assimilate children into reasonably responsible and tolerant city life." Social order in this vision does not require personal friendship or kinship ties but rather collective expectations for action in the public sphere.

of analysis is thus the neighborhood, and our phenomenon of interest is the physical and social disorder within its purview.[8]

Relatedly, the theoretical logic of collective efficacy focuses foremost on activity patterns that can be visibly observed. Here is where the theory links most naturally to the SSO method. Disputes among acquaintances and domestic violence that occurs indoors, for example, are by their very nature less amenable to public surveillance and sanctioning. They very likely are also perceived (albeit incorrectly) as less threatening to the common good. Collective efficacy is thus particularly relevant to explaining the incidence of crime and disorder in public spaces, and to crimes like robbery and burglary that typically elicit target selection decisions based on visual cues.[9] It follows that if the theory of collective efficacy is valid, it should be able to explain variations in independently collected measures of disorder in public spaces. We therefore introduce the method of systematic observation as a critical test case for the hypothesis that collective efficacy inhibits neighborhood disorder.

Finally, by implication, our theory offers a different way to think about the question of ecological "comorbidity"—the association between public disorder and predatory crime, especially violence. The "broken windows" literature sees disorder as a fundamental cause of crime (Skogan 1990, p. 75; Kelling and Coles 1996); if true, the hypothesized association of structural characteristics and collective efficacy with crime and violence ought to be largely mediated by social disorder. The alternative hypothesis we offer is that disorder is a manifestation of crime-relevant mechanisms and that collective efficacy should reduce disorder *and* violence by disempowering the forces that produce both. Our theory suggests also that structural constraints such as resource disadvantage and mixed land use account for both crime and disorder simultaneously. We thus test whether neighborhood disorder is an essential link in the ecological pathway that leads to predatory crime or rather a spuriously related construct rooted in structural constraints and collective social processes.

[8] Parenthetically, we believe that the routine activity insight has not sufficiently penetrated "neighborhood effects" research in sociology. A thought experiment reveals the possibility that 100% of the incidents of crime (or disorder) could be concentrated in one neighborhood (e.g., stimulated by opportunities and low surveillance) but with constancy in offending across neighborhoods (i.e., each contributing an equal rate of offenders). The former would suggest a neighborhood effect and the latter none. More generally, neighborhood research on behaviors such as academic achievement, employment, and drug use is problematic to the extent that the behaviors of interest are embedded in multiple ecological contexts (e.g., schools, parks, businesses) outside the respondent's neighborhood of residence. Clearly, a plausible theory of the ecology of the phenomenon under study is the first order of business.

[9] See Wright and Decker (1997). Note also that a surprising proportion of property crimes (almost 50%) occurs in public places (Felson 1987).

TABLE 1

Distribution of Sampled Chicago Census Tracts and Neighborhood Clusters

Race/Ethnicity	SES Low	SES Medium	SES High
75% black or more	31 (9)	10 (4)	9 (4)
75% white or more	0 (0)	7 (4)	18 (8)
75% latino or more	12 (4)	12 (4)	0 (0)
20% latino or more/20% white or more	11 (4)	14 (5)	10 (4)
20% latino or more/20% black or more	7 (4)	7 (4)	0 (0)
20% black or more/20% white or more	3 (2)	4 (4)	10 (4)
Other heterogeneous	8 (4)	14 (4)	9 (4)
Total	72 (27)	68 (29)	56 (24)

Note.—SES was defined by a six-item index that summed standardized census-based measures of median income, % college educated, % with household income over $50,000, % families below the poverty line, % on public assistance, and % with household income less than $5,000. The last three items were reverse coded. Neighborhood clusters are given in parentheses.

RESEARCH DESIGN

Between June and October 1995, observers trained at the National Opinion Research Center (NORC) drove a sport utility vehicle (SUV) at a rate of five miles per hour down every street in 196 Chicago census tracts. These tracts were selected from a stratified probability sample to maximize variation by race/ethnicity and SES. As part of a larger study, 343 "neighborhood clusters" (NCs) representing combinations of all 865 census tracts in Chicago were first stratified by seven categories of race/ethnic mix and three levels of SES.[10] Within strata, 80 NCs were then sampled for intensive study with the aim of obtaining a near balanced design, eliminating the confounding between ethnic mix and SES (table 1). However, in Chicago as in many other cities, ecological sorting by race and class results in a sample with some empty cells—low SES, predominantly European-American; high SES, predominantly Latino; and high SES, mixed Latino and black. Also, the largest stratum was low SES and predominantly African-American and contained 77 NCs, generally characterized by concentrated poverty, racial segregation, and other forms of disadvantage. The final design randomly sampled four NCs within cells that had

[10] Containing about 8,000 people, NCs are composed of geographically contiguous census tracts with similar distributions on key census indicators (e.g., race, SES, family structure, housing). We also used geographic boundaries (e.g., railroad tracks, parks, and freeways) and our knowledge of Chicago's neighborhoods as guides for constructing NCs. For more details, see Sampson et al. (1997, p. 924).

at least four, all NCs within cells having fewer than four, with an oversampling of the largest and most disadvantaged cell. Although reflecting the reality of segregation by race and class, the sampled NCs nonetheless tap the maximal existing variation across these dimensions of stratification. The distribution of the 196 census tracts embedded within the set of 80 sampled neighborhood clusters is shown in table 1.

The geographic unit of recorded observation within the sampled NCs and tracts was the face block: the block segment on one side of a street. For example, the buildings across the street from one another on any city block comprised two separate units of observation. At each intersection, a unique geographic identification code was assigned so that adjacent block faces could be pieced together to form higher levels of aggregation desired by theory or as suggested by patterns in the data. To observe each block face, the NORC team fielded a driver, a videographer, and two observers. As the SUV was driven down the street, a pair of video recorders, one located on each side of the SUV, captured social activities and physical features of both face blocks simultaneously. At the same time, two trained observers, one on each side of the SUV, recorded their observations onto an observer log for each block face. The observers added commentary when relevant (e.g., about unusual events such as an accident or drug bust) by speaking into the videotape audio. Face blocks were observed and videotaped between the hours of 7 A.M. and 7 P.M.[11] Applying these procedures, the SSO team produced videotapes, observer logs, and audiotapes for every face block in each of the 80 sampled NCs.

In all, 23,816 face blocks were observed and videotaped, for an average of 298 per NC and 120 per tract.[12] The data collected from the 23,816 observer logs focus mainly on land use, traffic, the physical condition of buildings, and evidence of physical disorder. Unlike the observer logs, which could be directly entered into machine-readable data files, the videotapes required the expensive and time-consuming task of first viewing and then coding. We thus selected a random subsample of all face blocks for coding. In those NCs consisting of 150 or fewer face blocks, all face blocks were coded; in the remaining NCs, sample sizes were selected to

[11] Although it would be desirable to assess disorder during the nighttime hours, a pretest confirmed that this was not feasible with current videotaping technology (or with the naked eye).

[12] The coding protocols for the observation logs and videotapes were first piloted in five census tracts from another city (Boston). After some modification, the protocols were then evaluated in a series of pretests in nonsampled neighborhoods of Chicago. Based on this experience, all major operational decisions were finalized. Coders were then trained in multiple sessions, including an intercoder reliability training where 90 face blocks were independently double coded, differences resolved, and coding procedures revised.

approximate a balanced design as closely as possible in order to maximize statistical power for comparisons of NCs and tracts. A total of 15,141 face blocks were sampled for videotape coding, an average of 189 per NC and 77 per tract. From the videotapes, 126 variables were coded, including detailed information on physical conditions, housing characteristics, businesses, and social interactions occurring on each face block (NORC 1995). As a check on quality control, new observers recoded a random 10% of all coded face blocks, and the results compared. This test produced over 98% agreement (NORC 1995; Carter, Dougherty, and Grigorian 1996).

Measures

Although some of the items were measured initially on an ordinal scale, the data behaved essentially as dichotomous items and so were coded for analysis as 1 = presence and 0 = absence of the indicator of disorder (Raudenbush and Sampson 1999*b*). The first scale is based on 10 items intended to capture the presence or absence of *physical* disorder. In declining order of observed frequency, the scale items include the presence or absence of *cigarettes or cigars* in the street or gutter (no = 6,815; yes = 16,758); *garbage or litter* on street or sidewalk (no = 11,680; yes = 11,925); *empty beer bottles* visible in the street (no = 17,653; yes = 5,870); *tagging graffiti* (no = 12,859; yes = 2,252); *graffiti painted over* (no = 13,390; yes = 1,721); *gang graffiti* (no = 14,138; yes = 973); *abandoned cars* (no = 22,782; yes = 806); *condoms* on sidewalk (no = 23,331; yes = 231); *needles/syringes* on sidewalk (no = 23,392; yes = 173); and *political message graffiti* (no = 15,097; yes = 14). The variation in sample size reflects the fact that six of the ten items were taken from the observer logs and thus have nearly complete data. The other four variables were derived from the videotapes and thus are based on the reduced subsample selected for coding. As expected, less serious indicators of disorder in public spaces (e.g., presence of cigarettes and garbage) arise more frequently than do indicators that might be regarded as more serious (e.g., drug paraphernalia and gang graffiti), with the presence of beer bottles arising with moderate frequency. Political graffiti was very rarely observed, even though it is not necessarily an indicator of severe disorder.[13]

[13] Detailed memos on the recognition of graffiti and gang insignia guided the training of coders (NORC 1995). Graffiti was classified by type—tagging, gang, and political—based on investigator guidelines, gang research in Chicago (Spergel 1995), and internal Chicago police memos on gang identification. Tag graffiti is identified by stylized forms such as block letter art or by attempts to create some form of visual expression. Gang graffiti is ordinarily distinguished by the absence of tag art and usually includes messages that refer to the name of a rival gang. Gang symbols in Chicago are longstanding and widely recognized and typically involve a combination of stars, crowns, emblems, and specific colors that distinguish among gangs (Spergel 1995; NORC

Direct evidence of *social* disorder was coded from the videotapes. The scale items include presence or absence of *adults loitering or congregating*[14] (no = 14,250; yes = 861); *drinking alcohol in public* (no = 15,075; yes = 36); *peer group with gang indicators* present (no = 15,091; yes = 20); *public intoxication* (no = 15,093; yes = 18); *adults fighting* or arguing in a hostile manner (no = 15,099; yes = 12); *selling drugs* (no = 15,099; yes = 12); and *prostitutes* on the street (no = 15,100; yes = 11). Most items in the disorder scale bear a conceptual affinity with concurrent "crime" in the sense of violation. Moreover, the frequency distribution tells us that, like crime, social disorder is quite rare—at least during the daylight hours. If social disorder resembles crime, indicators of the former should also be present far less frequently than indicators of physical disorder. The frequencies support this inference too. Activities that might be viewed as indicating more serious disorder (prostitution, drug selling, adults fighting) are especially rare. Indicators that are somewhat less serious are also relatively less rare (drinking alcohol, presence of peer gangs), though they remain rare overall relative to physical disorder. One item—adults loitering—occurred with much higher frequency than did any other item.

The frequency distribution of items suggests that the physical disorder scale will behave better as an ecological measure than will the social disorder scale. It not only has more items (10 vs. 7) but more important, the physical disorder items range widely; several occur with large frequency, several others with modest frequency, and several are comparatively rare. By contrast, the social disorder indicators occur with rare frequency except for adults loitering or congregating. The social disorder scale may thus be dominated by a single item of relatively low frequency, leading to an overall lack of between-neighborhood reliability.

Raudenbush and Sampson (1999*b*) adapted tools found useful in psychometrics to the problem of evaluating the systematic social observation of disorder. There are multiple components of measurement error that they addressed: (*a*) item inconsistency within a face block, (*b*) face-block variation within larger geographic units, and (*c*) temporal variation. The latter is a particular problem in measuring social disorder (Perkins et al. 1992; Taylor 1999). The probability of finding adults loitering or drinking, of finding peer gangs hanging out, or of seeing prostitution or drug deals

1995). Political graffiti was defined to include political messages such as "Stop the war in El Salvador" or "Reelect Mayor Daley."

[14] Because children and teenagers commonly play in public spaces, the coding rules limited the observation of loitering to groups of three or more adults not waiting for scheduled activities or businesses. For example, groups of adults waiting for public transportation or in line for a store would not be included. Coders were trained to a high degree of interrater agreement across all types of neighborhoods. Still, we acknowledge some residual ambiguity in the meaning of "loitering."

will depend on the time of day in which a face block is observed. Thus it is necessary to estimate and adjust for time of day. Fortunately, because face blocks were assessed between the hours of 7 A.M. and 7 P.M., there is considerable temporal variation in the time of observation within each census tract. We also need to allow for randomly missing data, because only a random sample of face blocks yielded data coded from videotapes.

To achieve these goals, we modify the three-level hierarchical regression model described by Raudenbush and Sampson (1999*b*).[15] The level-1 units are scale-item responses within face blocks, the level-2 units are face blocks, and the level-3 units are the 196 census tracts embedded within the sampled 80 NCs. In this article, we operationalize neighborhood using the 196 census tracts rather than NCs for three important reasons. First, previous research has argued for the smallest level of aggregation possible in measuring observed disorder, owing to considerable variability block-to-block within larger ecological units (Perkins et al. 1992; Taylor 1997). Second, by dropping down to the tract level from NCs, we more than double the neighborhood-level degrees of freedom, thus providing more statistical power to detect between-area variations. The number of face blocks observed within each tract was large enough to produce reliable measures of mean tract-level differences; the same was *not* true for block groups or block-level measures. Third, census tracts provide the additional information necessary to address the well-known multicollinearity among ecological variables (Land et al. 1990). At the tract level, the overlap among variables is considerably less than for NCs.

The appendix presents the three-level statistical model and measurement properties of the physical disorder and social disorder scales that are the main substantive focus of the ensuing analysis. The results show that the two disorder scales are highly reliable at the ecological level. Differences between neighborhoods in their aggregated disorder scores can be interpreted as expected differences in the log-odds of finding disorder on a typical item in the scale, adjusted for time-of-day effects in observation. The scales used below are thus meaningfully interpretable and arguably a linear (interval) scale appropriate for analysis via standard linear models.

INDEPENDENT DATA COLLECTION

To assess the theoretical framework on collective efficacy, structural constraints, and observed public disorder, we examine independent sources of data collected from a survey, police records, vital statistics, and the

[15] The initial methodological work for the SSO model was conducted on the 80 NCs, generally with similar measurement results (for details, see Raudenbush and Sampson 1999*b*).

census. First, over 4,000 households within the 196 tracts and 80 NCs were selected according to a multistage probability sample (see Sampson et al. 1997 for details). Within each household, a randomly chosen adult was interviewed in late 1994 and in the first eight months of 1995 concerning conditions and social relationships in the local neighborhood. The final sample size was 3,864, reflecting a response rate of 78%.[16]

Derived from the theoretical strategy outlined earlier, we examine four constructs from the survey—disorder, cohesion, control, and crime—aggregated to the tract level. Predatory crime was measured from respondents' reports of whether they (or any member of the household) had experienced within the past 6 months (*a*) a violent victimization in the neighborhood or (*b*) a household burglary or theft victimization. Approximately 5% of respondents reported a violent crime and 16% a household crime. The multi-item disorder scale taps how much a problem ("a big problem," "somewhat of a problem," "not a problem") residents rated the presence in the neighborhood of social incivilities (drinking in public, selling or using drugs, teenagers causing a disturbance) and physical incivilities (litter, graffiti, vacant housing). The reliability of the combined scale of perceived disorder at the tract level is .83 (for details on aggregate-level reliability see Raudenbush and Sampson 1999*a*).

A measure of shared expectations for informal social control was represented by a five-item Likert-type scale. Residents were asked about the likelihood ("Would you say it is very likely, likely, neither likely nor unlikely, unlikely, or very unlikely?") that their neighbors could be counted on to take action ("do something") if (*a*) children were skipping school and hanging out on a street corner, (*b*) children were spray painting graffiti on a local building, (*c*) children were showing disrespect to an adult, (*d*) a fight broke out in front of their house, and (*e*) the fire station closest to home was threatened with budget cuts. "Social cohesion/trust" was represented by five conceptually related items. Respondents were asked how strongly they agreed (on a 5-point scale) that "People around here are willing to help their neighbors," "This is a close-knit neighborhood," "People in this neighborhood can be trusted," and, reverse coded, "People in this neighborhood generally don't get along with each other" and "People in this neighborhood do not share the same values." Social cohesion and informal social control were correlated at $r = .68$ ($P < .01$), suggesting that the collective willingness to intervene in the neighborhood is enhanced under conditions of mutual trust and cohesion. Following Sampson et al. (1997), the two measures were combined to create a more parsi-

[16] The average number of survey respondents per tract was 20. One primarily nonresidential tract was not available for further analysis because it had only one survey respondent (with missing data).

monious and readily interpretable measure of "collective efficacy" with strong theoretical connections to disorder reduction. The aggregate reliability is .68 and .80 at the tract and NC levels, respectively.

The second set of data taps ecological variations in crime independent of the survey. The main source consists of incidents (not arrests) of homicide, robbery, and burglary in the years 1993 and 1995, geocoded from records of the Chicago Police Department and aggregated to the census tract of occurrence.[17] To provide some purchase on controlling for prior sources of crime not captured in our measured variables, 1993 data are employed. The crimes of homicide, robbery, and burglary are generally well reported and provide insight on variation in predatory crimes involving both property and persons. Comparatively, however, we place the most confidence in homicide as it is the most reliably measured of all crimes and does not suffer major reporting limitations. Because homicide is so rare—50% of the tracts had no incidents—we analyze the raw count of homicide using a negative binomial regression.[18] By contrast, all neighborhoods had recorded incidents of burglary and robbery, with a mean number around 40. We analyze the log of the robbery rate per 100,000 persons (mean = 6.90) and the log of the burglary rate per 100,000 households (mean = 8.32) as outcomes, both of which approximate very well a normal distribution. Overall, this research strategy provides five indicators of crime-rate variation covering property and violent crimes, and as measured in both surveys and police incident data. In addition, a person-based measure of homicide victimization derived from vital statistics rather than police records is employed in later analysis (described below).

The third source of independent data was culled from 1990 census data at the tract level. Three indexes of neighborhood structural differentiation are examined based on prior theory (Wilson 1987; Sampson et al. 1997) and research analyzing census data in Chicago over three decades (Morenoff and Sampson 1997). To reduce the dimensionality of the data, an alpha-scoring factor analysis with an oblique factor rotation was performed on the tract level, replicating the results of Sampson et al. (1997, p. 920) at the NC level. *Concentrated disadvantage* represents an economic disadvantage factor in racially segregated urban neighborhoods that was dominated by high loadings (> .8) for poverty, public assistance, unemployment, and female-headed families. Percentage of black residents was also linked to this dimension, although to a lesser extent (.62). Hence, this factor reflects the neighborhood concentration of resource disadvantage,

[17] We owe a debt of gratitude to Richard Block for providing the raw data.

[18] Poisson regression yielded evidence of overdispersion and thus the desirability of a negative binomial model, which incorporates a variance parameter to represent heterogeneity across tracts.

to which African-Americans and single-parent families with children are disproportionately exposed (Wilson 1987; Land et al. 1990). The second factor captures areas of *concentrated immigration.* The variables that loaded high on this dimension were percentage Latino (.95), percentage foreign-born (.73), and, to a lesser extent, density of children (percentage of persons ages 6–15), which loaded at .6. The third factor was dominated by two variables with very high (> .8) loadings—percentage living in the same house as five years earlier, and percentage of owner-occupied homes. The emergence of a *residential stability* factor is consistent with much past research (Kasarda and Janowitz 1974). Using factor loadings as weights, summary scales were created to reflect the three dimensions.[19]

As noted earlier, the routine activities approach (Cohen and Felson 1979; Cohen et al. 1981; Wikström 1991) suggests that ecological characteristics of neighborhoods reflect opportunities for crime and bear on the ability of residents to engage in guardianship. We therefore control for two ecological constructs emphasized in this approach—land use and density. Using census data, we control for the number of persons per square kilometer in the tract (mean = 7,530). Neighborhoods with more people per unit of space may generate greater anonymity and persons in public, making it harder for residents to maintain informal social control over public space.[20] The second control for neighborhood ecology is land use, defined as the proportion of face blocks in the tract that contain mixed residential and commercial activity (mean = 25%). Mixed land use has been shown to be a robust but understudied correlate of crime and disorder (see, e.g., Wikström 1991; Taylor 1995) and is theoretically relevant to understanding collective efficacy as well. It may be, for example, that the capacity of residents to achieve common purpose is limited not because of lack of internal effort but simply the structural constraint imposed by the density of commercial traffic and land-use patterns inhospitable to social interaction and surveillance. The mixed land-use variable is well measured, constructed from the complete set of 23,816 observational coding logs.

SOURCES AND CONSEQUENCES OF DISORDER

Observed disorder is correlated at the tract level with those constructs measured in the community survey and other independent sources (census

[19] A principal components analysis produced equivalent results (e.g., the tract-level correlation between a principal components and oblique rotation for the poverty scale was .99).

[20] We examined several other indicators of density and street activity, including housing density and the number of workers taking public transportation per capita. However, these were either highly correlated with population density or added little to the explanatory models.

data, official police records) most theoretically linked to it. For example, a healthy Pearson correlation of $r = .56$ ($P < .01$) emerges between SSO social disorder and social disorder measured in the community survey. SSO physical disorder is correlated .55 ($P < .01$) with survey-reported physical disorder. It is noteworthy that the correlation between residents' perceptions of physical disorder and SSO physical disorder is nearly identical to the correlation between survey-reported social disorder and SSO social disorder, given that physical disorder is relatively stable over time and social disorder reflects events that occur randomly in time. SSO disorder is also moderately strongly correlated with the survey measures of collective efficacy ($r = -.49$ for social disorder; $r = -.47$ for physical disorder; $P < .01$) in the direction expected.

Turning next to correlations with sociodemographic composition, SSO physical disorder is significantly related to census measures of concentrated poverty ($r = .50$) and immigrant concentration ($r = .39$). The relationship with residential stability is weak and insignificant. Furthermore, physical disorder measured in the SSO is only moderately correlated with rates of predatory crime measured by police-recorded rates of homicide ($r = .27$), robbery ($r = .45$), and burglary ($r = .24$). The relationships with survey-reported victimization are weaker: $r = .21$ for violence, and $r = .06$, NS, for burglary. A similar pattern of correlation appears with respect to social disorder. The bivariate results thus suggest that SSO measures of disorder have reasonably consistent relationships with theoretically linked explanatory factors derived from the neighborhood survey and census. Notably, however, the correlations of SSO disorder with crime rates, although positive, are not at the levels one might expect from the broken windows thesis.

Table 2 begins to unpack the sources and consequences of observed disorder across our sample of 195 Chicago census tracts in theoretically specified multivariate models. We examine first the weighted least squares (WLS)[21] regression of physical and social disorder measured by SSO on dimensions of structural differentiation, collective efficacy, density, and land use.[22] With approximately 50% of the variance explained, note the

[21] As noted above, the disorder scales form a linear and near-normally distributed metric. However, because the number of blocks used to create the scales varied by tract, WLS regression was used to induce homoscedasticity of error variances. Each case was weighted by the square root of the unweighted number of assessed blocks, giving more weight to tracts with a larger sample of coded data (see Hanushek and Jackson 1977, pp. 143, 152).

[22] Using the three-level hierarchical logistic regression model for uncertainty in measuring observed disorder elucidated in the appendix, the tract-level measures account for the multiple sources of measurement error identified (e.g., temporal, missing data). In this framework, empirical Bayes (EB) residuals are analyzed, defined as the least-squares residuals regressed toward zero by a factor proportional to their unreliability

TABLE 2

WLS Tract-Level Regression of SSO Physical and Social Disorder on Census Structural Characteristics, SSO Land Use, and Survey-Based Collective Efficacy

	Systematic Social Observation			
	Physical Disorder		Social Disorder	
	B	*t*-ratio	B	*t*-ratio
Concentrated disadvantage	.44	7.56**	.40	6.58**
Residential stability	.01	.12	−.02	−.36
Immigrant concentration	.28	4.99**	.19	3.20**
Population density	.03	.57	.01	.11
Mixed land use	.18	3.19**	.24	4.08**
Collective efficacy	−.19	−2.54*	−.21	−2.76**
Adjusted R^2	.52		.48	

Note.—$N = 195$.
* $P < .05$.
** $P < .01$.

strong association of disorder with concentrated disadvantage and immigrant concentration. The standardized coefficients for concentrated disadvantage are by far the largest in predicting physical disorder (B = .44; *t*-ratio = 7.56) and social disorder (B = .40; *t*-ratio = 6.58). After accounting for structural aspects of neighborhood differentiation by class and race/ethnicity, we also observe significant relationships with land use. Neighborhoods with mixed residential and commercial development exhibit higher levels of both physical and social disorder, regardless of sociodemographic characteristics. Furthermore, the estimated association of collective efficacy with both forms of disorder is consistently negative and significant.

The data thus far suggest that structural constraints matter greatly in predicting disorder. Moreover, collective efficacy with respect to the social control of public space appears to inhibit the incidence of observed disorder, whether physical or social in nature. That this relationship is somewhat stronger with social disorder (unstandardized b = −.81) than physical disorder (b = −.65) is consistent with the theoretical underpinnings

(Bryk and Raudenbush 1992, chap. 3). Similarly, the collective efficacy and perceived disorder scores for each tract were computed using the EB residuals, also accounting for measurement error and missing data (see also Raudenbush and Sampson 1999*a*). Using EB residuals as explanatory variables corrects for bias in regression coefficients resulting from measurement error (Whittemore 1989). Although more precise, the EB results nonetheless mirror the results using simple scale averages.

TABLE 3

WLS Tract-Level Regression of SSO Physical and Social Disorder on Prior Homicide Rate, Perceived Disorder, and Collective Efficacy

	Systematic Social Observation			
	Physical Disorder		Social Disorder	
	B	t-ratio	B	t-ratio
Survey physical disorder	.41	5.89**	. . .	
Survey social disorder	. . .		.41	5.36**
Prior (1993) homicide	.15	2.59*	.08	1.39
Collective efficacy	−.26	−3.69**	−.25	−3.28**
Adjusted R^2	.42		.39	

Note.—$N = 195$.
* $P < .05$.
** $P < .01$.

of collective efficacy. Clearly, however, concentrated disadvantage is the single most important predictor of disorder in Chicago neighborhoods. The structural origins of disorder cannot be ignored (Hunter 1985).[23]

Table 3 presents a methodologically oriented test of the association of collective efficacy with observed disorder. It may be that respondents' perceptions of disorder in the neighborhood colored their judgments about cohesion and control. That is, informants might reasonably infer low collective efficacy from their perceptions of disorder in the neighborhood. Because survey disorder (combined) is correlated −.63 with collective efficacy and .62 with SSO disorder ($P < .01$), controlling its effects serves as a strict test. After all, in order for disorder to confound the efficacy scale, it must be mediated by the perceptions of the reporter. Accordingly, we estimate the net effect of efficacy on disorder *after* resident perceptions of disorder have been removed. Because the vast majority of the survey was completed before the SSO, this control might also be interpreted as estimating the effects of collective efficacy on changes in disorder. To provide another guard against the endogeneity of collective efficacy, we control for prior violence as measured by the 1993 homicide rate. Total vio-

[23] All regression analyses were examined for influential observations using a variety of diagnostics (e.g., Cook's D; leverage scores). No one tract exerted a disproportionate influence on the coefficient estimates after appropriate (e.g., log) transformations. We also examined various diagnostics for multicollinearity. At the tract level, this did not turn out to be a problem. In fact, all correlations among predictor variables were less than .50, and the variance inflation factors (VIF) were all less than 2.5, well below those levels traditionally thought to be of concern (see Belsley, Kuh, and Welsch 1980).

lence showed similar but weaker results. Thus, both lagged homicide and concurrent perceived disorder are used to adjust the levels of observed disorder, disaggregated in both the survey and SSO measures by type.

The data in table 3 are very clear in showing a consistent negative relationship of collective efficacy with SSO disorder, both physical and social. The estimated net effect of collective efficacy on physical disorder is $B = -.26$ (t-ratio $= -3.69$), larger than the (significant) direct effect of prior homicide ($B = .15$). The effect of collective efficacy on social disorder is the same magnitude ($B = -.25$; $P < .01$). Not surprisingly, the direct association of perceived with observed disorder is larger, regardless of whether physical or social in nature. The key result, however, is that whether we control for sociodemographic characteristics from the census, prior violence, or even perceived disorder, the data show a persistent negative association of collective efficacy with the independently observed incidence of disorder.[24] Because the results in tables 2–3 are so similar for physical and social disorder, and because the two scales are highly correlated ($r = .71$ for EB residuals), we combine them into a summary index of SSO disorder for the remaining analysis.

"Broken Windows" Revisited

In the next set of tests, we turn to the consequences of disorder. A fundamental thesis of "broken windows" is that observed disorder directly causes predatory or "serious" crime. An alternative interpretation is that disorder and crime are both the products of weakened social controls and structural antecedents. Table 4 provides evidence that adjudicates between these two scenarios using survey-reported victimization aggregated to the neighborhood level. Model 1 in the top half of the table, with no controls for structural characteristics or collective efficacy, shows that SSO disorder *is* positively linked to survey-reported violence (t-ratio $=$ 2.50; $P < .05$), although only 5% of the overall variance is explained. Interestingly, officially measured rates of prior violence are not significantly related to later survey-reported violence. By contrast, 1993 burglary rates predict later survey-reported burglary, while SSO disorder is seen to be unimportant. This reveals that disorder, systematically measured with observation, is positively but rather weakly related to survey-reported neighborhood victimization.

[24] When the census structural factors are added to table 3, the significant link between collective efficacy and SSO disorder is maintained. By contrast, the survey-based measure of disorder, which is correlated at $r = .63$ ($P < .01$) with concentrated disadvantage, is rendered insignificant. Unlike collective efficacy, then, perceptions of disorder appear to be confounded with population composition.

TABLE 4

WLS Tract-Level Regression of Survey-Reported Victimization in the Neighborhood

	Survey-Reported Victimization, 1995			
	Personal Violence		Household Burglary	
	B	t-ratio	B	t-ratio
Model 1:				
SSO disorder	.20	2.50*	.08	1.10
Prior (1993) crime	.08	1.03	.22	3.16**
Adjusted R^2	.05		.05	
Model 2:				
SSO disorder	.12	1.11	−.02	−.16
Prior (1993) crime	.00	.02	.24	3.15**
Concentrated disadvantage	−.03	−.25	−.14	−1.51
Residential stability	−.04	−.44	.07	.74
Immigrant concentration	.10	1.16	.23	2.75**
Population density	−.16	−1.65	−.14	−1.53
Mixed land use	−.17	−2.05*	.04	.45
Collective efficacy	−.37	−3.59**	−.22	−2.16*
Adjusted R^2	.13		.13	

Note.—N = 195. For 1995 survey violence, the control for prior crime is the logged incident rate in 1993 of murder, rape, robbery, and aggravated assault per capita. For burglary victimization, the control is the logged rate of 1993 burglaries per household.

* $P < .05$.

** $P < .01$.

In the bottom half of table 4 (model 2), we reestimate the effects of disorder and prior crime after adding controls for structural antecedents and collective efficacy. The results are consistent and point to a spurious association of disorder with predatory crime. In the case of violent victimization, the coefficient for SSO disorder is cut appreciably and reduced to insignificance. Collective efficacy, on the other hand, is by far the largest predictor of violent victimization (B = −.37; t-ratio = −3.59), absorbing the prior effects of concentrated poverty, residential stability, and immigrant concentration (see also Sampson et al. 1997). In fact, controlling for collective efficacy alone eliminated the effect of disorder. It is noteworthy as well that collective efficacy is associated with lower rates of violence regardless of sociodemographic composition and crime-linked mechanisms—namely, prior violence and observed neighborhood disorder. Columns 3 and 4 reveal a similar picture for household burglary victimization. Here we see that collective efficacy is linked to significantly lower burglary rates (B = −.22). Latino neighborhoods undergoing immigration flows experience higher rates of survey-reported household victimization.

SSO disorder showed almost no explanatory power in the first place (t-ratio = 1.10), and its coefficient in the full model is reduced to near zero.

We conducted several other tests to assess the robustness of the results for survey-reported victimization. Because violent victimization is rare, the measure was highly skewed—some 50% of the areas yielded no incidents. We therefore re-estimated the violent victimization model using logistic regression, with neighborhoods experiencing no incidents coded "0" and neighborhoods with one or more coded "1". The estimated effect of collective efficacy was still the largest (t-ratio = −3.22), and SSO disorder remained insignificant (t-ratio = .83). We also examined a summary measure of total reported victimization in the neighborhood, and in addition, we addressed shared method variance by controlling for perceived disorder. Reported victimization is correlated significantly with perceived disorder and may thereby confound informant reports of neighborhood collective efficacy. Under this specification, the effect of collective efficacy remained significant (t-ratio = −3.79), while the SSO-disorder link remained insignificant. Collective efficacy also predicted survey-reported violence (B = −.39) controlling for perceived disorder, whereas perceived disorder exhibited a null effect. Finally, we examined SSO physical disorder and social disorder in separate models to see if there were differential effects; neither measure emerged as a significant predictor. To this point then, it cannot reasonably be concluded that disorder is a proximate or strong mechanism explaining crime-rate variation. Rather, the results suggest a direct association of collective efficacy with lower rates of predatory victimization, unmediated by social disorder.

Table 5 turns to the other set of independently measured outcomes—officially recorded incidents of homicide, robbery, and burglary in 1995. At the bivariate level (model 1), SSO disorder predicts all three types of official crime, especially robbery rates (B = .48; t-ratio = 7.67). Structural antecedents, land use, and collective efficacy are introduced next as covariates at the multivariate level (model 2). For burglary, we see that SSO disorder drops out of the picture. In predicting homicide, the estimated coefficient for disorder also drops (by half) but retains significance. Only for robbery does the estimated effect of disorder remain substantively large, with a standardized coefficient of .31 (t-ratio = 3.86). Across all three crimes, the consistent predictors are concentrated poverty and collective efficacy. The latter's coefficient for homicide is −1.97 (t-ratio = −4.12); further calculation indicates that a standard deviation increase in collective efficacy was associated with a 35% decrease in the expected homicide rate after adjusting for prior homicide, disorder, and structural antecedents. A similar outcome obtains for collective efficacy in predicting rates of robbery and burglary. In fact, collective efficacy has the largest standardized effect estimate on robbery (B = −.35) and the second largest

TABLE 5

TRACT-LEVEL REGRESSION OF POLICE-RECORDED CRIME INCIDENTS IN 1995

	Police-Recorded Incidents, 1995					
	Homicide Counts		Log Robbery Rate		Log Burglary Rate	
	Coefficient	z-ratio	B	t-ratio	B	t-ratio
Model 1:						
SSO disorder	.27	5.23**	.48	7.67**	.26	3.79**
Model 2:						
SSO disorder	.13	2.17*	.31	3.86**	.07	.75
Concentrated disadvantage	.27	2.80**	.26	3.59**	.23	2.71**
Residential stability	.24	2.27*	−.12	−1.74	.11	1.40
Immigrant concentration	−.14	−1.56	−.22	−3.51**	.12	1.52
Population density	−.00	−1.82	−.35	−5.10**	−.43	−5.30**
Mixed land use	.00	.12	.03	.51	.08	1.06
Collective efficacy	−1.97	−4.12**	−.35	−4.43**	−.27	−2.92**
R^2	.19		.46		.26	
Model 3:[a]						
SSO disorder	.09	1.56	.10	2.12*	.02	.29
Collective efficacy	−1.74	−3.51**	−.11	−2.38*	−.20	−2.63**
Prior (1993) crime	.26	2.02*	.83	21.01**	.59	10.15**
R^2	.20		.85		.52	

NOTE.—$N = 195$. R^2 for homicide counts are pseudo R^2s; for log robbery rate and log burglary rate, they are adjusted R^2s. Homicide events are analyzed with negative binomial regression, with log population as the exposure (control) variable (coefficient not shown); robbery and burglary refer to logged incident rates per 100,000 (persons for robbery, households for burglary) analyzed with WLS regression using square root of population as weight. For the 1995 homicide, robbery, and burglary equations, the control for prior crime refers to the 1993 log homicide count, robbery rate, and burglary rate, respectively.

[a] All structural characteristics shown above are controlled (coefficients not shown).

* $P < .05$.

** $P < .01$.

on burglary (B = −.27). For present purposes, then, the key result is that the influences of structural characteristics and collective efficacy on burglary, robbery, and homicide are not mediated by neighborhood disorder.

As in table 4, we now exploit the longitudinal nature of the police data and provide a further test of collective efficacy and disorder by adding a control for the lagged crime rate (model 3). Unlike table 4, this specification turns out to be a severe one for official robbery and burglary because of the very strong connection between rates in 1993 and 1995, a result we suspect arises in part because of the common measurement source. For example, the 1993 log robbery rate is correlated .92 with the log robbery rate in 1995. This level of serial dependence makes it difficult for *any* concurrent factor to wield much direct influence, providing overly conservative estimates.[25] Nevertheless, collective efficacy is directly related to lower rates of homicide, robbery, and burglary in 1995 after adjusting for 1993 levels of those crimes, as shown in the bottom portion of table 5. The coefficient for homicide (-1.74) remains large in magnitude, and the estimated effect for robbery (t-ratio $= -2.38$) is noteworthy in light of the very strong ($> .9$) positive association between 1993 and 1995 robbery coupled with the significant negative association between prior robbery and collective efficacy ($r = -.39$). By contrast, the association of disorder with homicide and burglary is reduced even further and is rendered insignificant for homicide.

The exception to the emerging conclusion that disorder is spuriously related to predatory crime is robbery. Columns 3–4 show that the effect of disorder is reduced but not eliminated by the introduction of a control for robbery in 1993. Areas with greater cues of disorder appear to be more attractive targets for robbery offenders, perhaps because disorder increases the potential pool of victims without full recourse to police protection, such as those involved in drug trafficking and prostitution. Wright and Decker's (1997) research has indicated that robbery offenders are especially attuned to local drug markets, where they perceive drug dealers and their customers as prime targets with cash on hand.[26]

Reciprocal Feedback

We close by addressing a concern that the estimated effect of collective efficacy on crime reflects reverse causation. It may be that neighborhood social trust and residents' sense of control are simultaneously undermined by crime, most notably interpersonal crimes of violence and those committed in public by strangers (Skogan 1990). Liska and Bellair's (1995) findings indicated that violent crimes such as robbery induce out-migration

[25] Note also that if collective efficacy has an insignificant direct effect in such a specification, this does not necessarily mean it is unimportant. Prior (unmeasured) levels of collective efficacy may have explained variation in prior crime, and thus the relationship with concurrent crime would be mediated.

[26] The three regressions of official crime rates were repeated with perceived disorder controlled; the main results were unchanged. The burglary and robbery models were also re-estimated using a negative binomial regression of raw counts, with logged population and logged counts of prior burglary and robbery controlled. The significant negative effect of collective efficacy on robbery and burglary was maintained, whereas the estimated effect of observed disorder on burglary was insignificant. Specifically, the negative binomial t-ratios reflecting the effect of collective efficacy on robbery and burglary events—with perceived disorder controlled—were -2.95 and -2.57, respectively. Further tests also revealed that physical disorder was somewhat more strongly related to robbery rates than social disorder when examined separately, presumably because physical disorder is the more reliable measure (see the appendix).

from central cities. Perhaps more relevant, Liska and Warner (1991) found that robbery constrains social interactions in public settings, thereby potentially dampening social cohesion and the emergence of shared expectations among residents for taking action to protect the community. Fear of crime, especially the fear of being accosted by strangers in public and attacked, may thus undermine neighborhood collective efficacy. By not accounting for the potential reciprocal effects of violence, we may have misestimated the role of collective efficacy relative to disorder. Our strategy to address this possibility has been to control for the prior incidence of violent events. Although providing a strict test, this strategy leaves unresolved the potential simultaneous relationship between collective efficacy and street crime. Moreover, because of the strong temporal dependence in the police record data, this strategy may have yielded unduly conservative estimates of both collective efficacy *and* disorder. We therefore address these issues by estimating a simultaneous equation model of violent crime and collective efficacy, and by introducing a new source of data on prior violence that is independent of police records.[27]

It is well known that estimating simultaneous relationships requires a priori identification restrictions that are, in practice, difficult to meet (Fisher and Nagin 1978). With respect to the issue at hand, however, we believe that extant theory provides reasonable grounds for specifying a causal feedback loop between violent crime and collective efficacy. First, we introduce measures of local social exchange, local friend/kinship ties, and neighborhood attachment as instrumental variables to identify the unique effect of collective efficacy.[28] A long line of urban research suggests that participation in social exchange, friend/kinship ties, and affective identification with the local area increases mutual trust and shared expectations for collective action in support of the neighborhood (Kasarda and Janowitz 1974; Sampson 1988). In the present case, we assume that friend/kinship ties, exchange, and affective sentiment for the neighborhood influence crime only insofar as they foster the core ingredients of collective efficacy. This seems reasonable, for there is little reason to expect that neighborly social ties or sentiment alone will reduce crime other than by

[27] We thank the reviewers for stressing the possibility of reciprocal causation, prompting us to push the analysis in new directions.

[28] The *local exchange* scale is based on five items in the survey that asked how often the respondent and neighbors exchange favors or goods (e.g., tools), have parties together, visit in each others' homes, watch each others' homes, and exchange advice or information. The measure of *friend/kinship ties* is derived from questions on the number of friends and relatives that respondents reported living in the neighborhood. *Attachment* is a two-item scale derived from questions on how sorry the respondent would be to move from the neighborhood and how satisfied he/she was with the neighborhood.

influencing the shared willingness to engage in informal social control under conditions of mutual trust (see also Bellair 1997).[29]

Second, to identify the crime equation, we created a new person-based index of the prevalence of prior homicide victimization among neighborhood residents, even if the incidents occurred elsewhere. The original source was death-record information (derived from the coroner's report) found in case-level vital statistics data for Chicago. Individual deaths by homicide victimization were geocoded and aggregated across years to yield more stable indicators. Based in part on a larger health-related project (Morenoff 1999), one indicator was constructed to reflect the age and sex-standardized rate of homicide victimization for the years 1989–91 for the larger NCs within which each tract is located. The second indicator is the logged rate of homicide deaths in each tract per 100,000 population for the years 1990–93. As expected, these two rates were positively correlated ($r = .62$; $P < .01$). To achieve parsimony, we combined them into a single index of resident-based homicide prevalence, which we use as an instrument to predict later incidents of violent crime.

Such a specification serves several theoretical purposes. It is well known in the criminological literature that victims and offenders in homicide transactions share similar demographic and residential profiles (Singer 1981). In fact, some 25% of homicides are victim precipitated (Wolfgang 1958). Thus, the prior rate of victim-based homicide is a proxy measure for resident offender rates of homicide production.[30] Moreover, it is plausible to assume that the homicide victimization rate of residents in 1989–93 will predict the later incidence of violence, but not directly affect levels of collective efficacy in 1995 independent of the concurrent rate of violent incidents. As Liska and Warner's (1991) research implies, it is the current presence of violence in their home neighborhoods—especially robbery—that residents most fear, not the past presence of victims, many of whom were victimized in other parts of the city. And, as noted above, there is evidence that residents are able to make the distinction between local

[29] Because local exchange and friend/kinship ties were rather highly correlated (.62; $P < .01$) and our goal is not to elucidate their independent effects but rather to identify the collective efficacy equation, for simplicity, we present the results for a combined scale of ties/exchange. We estimated models with each separately and found similar overall results.

[30] More specifically, this specification provides some purchase on assessing a "routine activity" model of violent events in the neighborhood while controlling for the differential composition of neighborhoods with respect to residents' involvement with violence. We acknowledge an anonymous reviewer for suggesting that we consider the violence potential of neighborhood residents. We sought to obtain offender residence data but were unsuccessful. The Chicago police do not release this information, and it is not available in vital statistics. Therefore we cannot directly assess offender production.

crime events and persons involved in criminal networks (Pattillo 1998). Therefore, we assume that any effect of resident-based homicide prevalence in 1989–93 on collective efficacy in 1995 works through its connection to the 1995 event rate of violence.

The maximum-likelihood estimates of the coefficients and t-ratios for the logged rate of homicide incidents per capita in 1995 are presented in table 6.[31] The fit of the model to the data is very good, with a chi-square of 9.86 relative to 6 degrees of freedom ($P > .10$) and an adjusted goodness of fit index of .90. The coefficients and standard errors in table 6 reveal that the estimated direct effects of the instrumental variables are substantial. Social ties/exchange and attachment each predict collective efficacy net of structural controls and homicide, while the direct effect of prior homicide victimization on the 1995 homicide rate is also significant (t-ratio = 2.85). The instrumental variables accounted for 22% of the unique variance in collective efficacy and 34% of the unique variance in homicide rates in the reduced-form equations. Most important, the fitted structural model indicates the presence of a reciprocal feedback between homicide and collective efficacy. Note the t-ratio of -2.20 for the estimated effect of collective efficacy on homicide and the t-ratio of -2.58 for the estimated effect of homicide on collective efficacy. Controlling for this feedback loop does not change our inference about the direct effect of social disorder on homicide. It remains insignificant (t-ratio = $-.38$).

Prior research on the fear of crime (Liska and Warner 1991; Liska and Bellair 1995) suggests that robbery might present an even more critical test, and our own data has shown that robbery is the only crime directly linked to social disorder (table 5). We thus estimated a simultaneous equation model using robbery instead of homicide as the measure of violence (not shown here, details available upon request). The results were in most respects similar to those in table 6. The instrumental variable for robbery was surprisingly effective, with prior homicide accounting for 30% of the unique variance in 1995 robbery.[32] And once again, there was a reciprocal association between collective efficacy and violence, with collective efficacy negatively related to robbery ($t = -2.40$) and robbery related nega-

[31] The models were estimated with LISREL ver. 8.04 (Joreskög and Sörbom 1993), with a weighted full-information, maximum-likelihood (FIML) procedure (unweighted results were similar). Recall that the logged-rate and event-count models produced similar results for both homicide and robbery (table 5). Also, recall that mixed land use consistently showed a large association with SSO disorder but not crime rates (see tables 4–5). Further analysis revealed an insignificant association of land use with collective efficacy. We therefore excluded land use in both the crime rate and collective efficacy equations, improving the overall model fit.

[32] The fit of the robbery model was not quite as good as for homicide, but it was still very adequate ($\chi^2 = 15.2$; $df = 6$; $P = .02$; AGFI = .85).

TABLE 6

MAXIMUM-LIKELIHOOD COEFFICIENTS AND t-RATIOS FOR SIMULTANEOUS EQUATION MODEL OF SSO DISORDER, COLLECTIVE EFFICACY, AND OFFICIAL HOMICIDE RATES

	ENDOGENOUS VARIABLES					
	Collective Efficacy[a]		SSO Disorder		Homicide Rate[a]	
	Coefficient	t-ratio	Coefficient	t-ratio	Coefficient	t-ratio
Exogenous:						
Concentrated disadvantage	−.04	−2.26*	.89	7.74**	−.04	−.17
Residential stability	.08	4.93**	−.02	−.17	.44	2.14*
Immigrant concentration	−.04	−3.61**	.51	4.64**	−.11	−.62
Population density	−.00	−2.83**	.00	.27	.00	.15
Mixed land use	NI		.04	4.19**	NI	
Ties/exchange[b]	.03	3.91**	NI		NI	
Attachment[b]	.11	3.25**	NI		NI	
Prior homicide[b]	NI		NI		.37	2.85**
Endogenous:						
Collective efficacy	NI		−1.69	−2.64**	−3.18	−2.20*
SSO disorder	NI		NI		−.03	−.38
Homicide rate	−.06	−2.58*	NI		NI	
R^2	.63		.58		.32	

NOTE.—N = 195. For overall fit, $\chi^2 = 9.86$; $df = 6$; and $P > .10$. NI = Not included in model specification.
[a] Error covariance for reciprocal relationship between collective efficacy and homicide = 8.72 (t-ratio = 2.24).
[b] Instrumental variable.
* $P < .05$.
** $P < .01$.

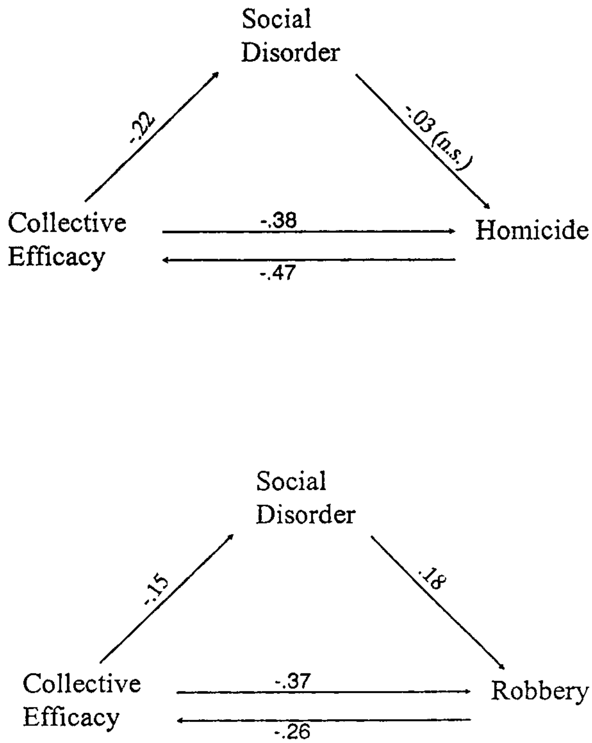

FIG. 1.—Structural model for relationships among endogenous variables—survey reported collective efficacy, observed social disorder, and official rates of violence. Coefficient estimates are from standardized solution and significant except where noted. For simplicity, exogenous variables are not displayed (see table 6 for a full listing of coefficients and t-ratios for homicide).

tively to collective efficacy ($t = -2.10$). Controlling for this reciprocal relationship, there remained a significant direct effect of social disorder on robbery (t-ratio $= 2.68$).

To aid in interpreting the pattern and magnitude of these results across crime types, standardized coefficients for the key structural relations of interest are shown in figure 1. The reciprocal relationship between collective efficacy and homicide is seen in the top half, with a collective efficacy to homicide path of $-.38$ and a homicide to collective efficacy path of $-.47$. The bottom half of figure 1 shows that robbery and collective efficacy are caught up in a similar feedback loop, net of structural characteristics, with a standardized coefficient of $-.37$ for the estimated effect of collective efficacy on robbery and $-.26$ for the reverse effect of robbery on collective efficacy. Interestingly, further comparisons reveal that for both robbery and homicide, the unstandardized coefficients for collective efficacy change very little in magnitude when the simultaneous relationship is modeled. For example, the coefficient estimate for collective efficacy on robbery in a recursive specification is -1.29 (SE = .29), compared to -1.33 (SE = .55) for the simultaneous model in the bottom half of figure 1. The corresponding coefficients for homicide are -3.54 (SE = .78) and -3.18 (SE = 1.57). Thus, the standard errors increase in the simultaneous equation results, but the substantive pattern remains the same (cf. tables 4–5).

The other major result obtained in figure 1 is that collective efficacy is inversely related to SSO disorder, which in turn exhibits a significant positive association with the robbery rate.[33] The standardized coefficient for SSO disorder on robbery is B = .18, similar in magnitude to the negative effect of collective efficacy on disorder (B = $-.15$). By contrast, the estimated effect of disorder on homicide is virtually zero. These patterns are consistent with earlier results from recursive models. To assess robustness, we re-estimated the models in figure 1 by specifying ties/exchange as the sole instrument for collective efficacy, and by separately introducing the constituent lagged indicators of homicide victimization as instruments for the current event rate. The results were substantively similar. We also controlled for the effects of perceived disorder on collective efficacy in both the robbery and homicide equations, specifying the former as a function of observed disorder. The results indicated that observed disorder increases perceived disorder, which in turn reduces collective efficacy. The significant reciprocal relationship between violence and collective efficacy nonetheless remained intact, as did the differential effect of observed disorder in predicting robbery but not homicide in figure 1.[34] The overall pattern in the data is thus clear. The estimated association of collective efficacy with disorder and crime holds up for a variety of measures and alternative theoretical specifications, including simultaneous causation, and thereby appears to be general in nature. By contrast, the direct disorder-crime association is specific to officially measured robbery.

[33] The estimated indirect effect of collective efficacy on robbery mediated by SSO disorder for the model in the bottom half of figure 1 is significant at $P < .10$ (t-ratio = -1.62).

[34] Furthermore, we estimated a model specifying a simultaneous relationship between collective efficacy and SSO disorder, using mixed land use as an instrument for disorder and, as in table 6, ties/exchange and attachment as instruments for collective efficacy. Yielding an excellent fit to the data ($\chi^2 = .005$; $df = 1$; $P = .94$), the model results indicated that the path from observed disorder to collective efficacy was insignificant, whereas the t-ratio for the estimated effect of collective efficacy on observed disorder was -2.74 ($P < .01$).

SUMMARY

The evidence presented in this study has shown that public disorder in urban spaces is a robust ecological construct that can be reliably measured at the neighborhood level using systematic observational procedures. Consistent with our theoretical expectations, structural characteristics—especially concentrated poverty and mixed land use—were strongly associated with physical and social disorder. Yet collective efficacy, defined as the fusion of social cohesion with shared expectations for the active social control of public space, predicted lower observed disorder after controlling not just sociodemographic and land-use characteristics, but perceived disorder and prior rates of predatory crime as well. Collective efficacy also maintained a significant relationship with violent crime after adjusting for simultaneous feedback effects.

On the other hand, observed disorder did not match the theoretical expectations set up by the main thesis of "broken windows" (Wilson and Kelling 1982; Kelling and Coles 1996). Disorder is a moderate correlate of predatory crime, and it varies consistently with antecedent neighborhood characteristics. Once these characteristics were taken into account, however, the connection between disorder and crime vanished in 4 out of 5 tests—including homicide, arguably our best measure of violence. The empirical results therefore support our contention that public disorder and most predatory crimes share similar theoretical features and are consequently explained by the same constructs at the neighborhood level, in particular the concentration of disadvantage and lowered collective efficacy.

Although our results contradict the strong version of the broken windows thesis, they do not imply the theoretical irrelevance of disorder. After all, our theoretical framework rests on the notion that physical and social disorder comprise highly visible cues to which neighborhood observers respond (see also Jacobs 1961; Goffman 1963; Lofland 1973; Skogan 1990; Taylor 1997). According to this view, disorder may turn out to be important for understanding migration patterns, investment by businesses, and overall neighborhood viability. Thus, if disorder operates in a cascading fashion—encouraging people to move (increasing residential instability) or discouraging efforts at building collective responses—it would indirectly have an effect on crime. Moreover, our results established a significant albeit relatively modest association of disorder with officially measured robbery. Apparently, robbery not only constrains social interaction and thus reduces social control (Liska and Warner 1991), but potential robbery offenders respond to visual cues of social and physical disorder in the neighborhood. Our findings regarding robbery also suggest a complex feedback loop (bottom half of figure 1), whereby disorder entices robbery, which in turn undermines collective efficacy, leading over time to yet more disorder and ultimately robbery.

What we would claim, however, is that the current fascination in policy circles (see Kelling and Coles 1996; Kelling 1998) on cleaning up disorder through law enforcement techniques appears simplistic and largely misplaced, at least in terms of directly fighting crime. Eradicating disorder *may* indirectly reduce crime by stabilizing neighborhoods, but the direct link as formulated by proponents was not the predominate one in our study. What we found instead is that neighborhoods high in disorder do not have higher crime rates in general than neighborhoods low in disorder once collective efficacy and structural antecedents are held constant (tables 4–5). Crime and disorder are not even that highly correlated in the first place. Even for robbery, the aggregate-level correlation does not exceed .5.[35] In this sense, and bearing in mind the example of some European and American cities (e.g., Amsterdam, San Francisco) where visible street-level activity linked to prostitution, drug use, and panhandling does not necessarily translate into high rates of violence, public disorder may not be so "criminogenic" after all in certain neighborhood and social contexts (see also Whyte 1988, pp. 156–64). Put differently, the active ingredients in crime seem to be structural disadvantage and attenuated collective efficacy more so than disorder. Attacking public disorder through tough police tactics may thus be a politically popular but perhaps analytically weak strategy to reduce crime, mainly because such a strategy leaves the common origins of both, but especially the last, untouched. A more subtle approach suggested by this article would look to how informal but collective efforts among residents to stem disorder may provide unanticipated benefits for increasing collective efficacy (Skogan and Hartnett 1998), in the long run lowering crime.[36]

Of course, several limitations of our study warrant further consideration. First, although the relationships that emerged were consistent under multiple and strict tests, until we have a longitudinal profile of neighbor-

[35] We also set aside for future research the possibility that the robbery finding reflects an artifact of official data. Recall that none of the survey measures of violence were directly related to disorder. It is possible that citizen calls to the police or police accuracy in recording robberies is greater in areas perceived to be high in disorder. That there was no disorder link for official homicide or burglary is perhaps telling.

[36] Informally mobilizing a neighborhood "clean up" to reduce physical disorder, for example, might build collective efficacy through the formation of new social ties and by increasing local awareness of the mutual commitment of residents to the area. Such mobilization might also demonstrate to participants and observers alike that people in the neighborhood could be relied upon to maintain public order and social control. A demonstration of formal control through a police-led crackdown on disorder might be expected to generate a different response, unless it too was internally mobilized.

hood change, the causal direction of effects rests in large part on a priori theory. Understanding the dynamics of disorder is a complex challenge and will require a new generation of longitudinal research. Second, and relatedly, our simultaneous equation model must be viewed cautiously despite its apparent robustness. Definitive tests of feedback processes await future research and validation. Third, the SSO method is limited to the extent that inferences from the visual record were systematically biased in ways not anticipated by the investigative team. It is possible, for example, that the trained observers used physical cues of disorder to inform judgments about social disorder.[37] The SSO training tried hard to avoid this scenario, but the fact is that all observations require inferences. Fourth, we acknowledge that the large number of ecological units assessed means that many contextual features were probably missed. SSO is thus not meant to be a replacement for the close ethnographic observation of neighborhoods found in a long-standing and continuing tradition of research (e.g., Whyte 1943; Lofland 1973; Pattillo 1998; Carr 1998). Rather, the whole point of *systematic* social observation is to allow for the comparative analysis of variations across a large number of analytically defined ecological areas. Our goal has therefore been to strive for a systematic method that is replicable (see also King, Keohane, and Verba 1993) and that serves as a complement to comparative urban ethnography. A fifth limitation to the approach proposed in this article is that we have not yet taken into account spatial autocorrelation (Raudenbush and Sampson 1999*b*). Rather, tract-level neighborhoods have been treated as independent.[38] Ongoing work will build spatial associations into the types of models presented here.

IMPLICATIONS FOR RESEARCH DESIGN

This article is part of a larger effort to build a *social* science of ecological assessment by formulating new research designs and statistical methods to improve the quality of "ecometric" measures (see also Raudenbush and Sampson 1999*a,* 1999*b*). Whereas psychometric procedures for the study

[37] However, this possibility is undermined by the fact that the results were not sensitive to substitution of the physical disorder for the social disorder SSO measure.

[38] This is not an unreasonable assumption for the SSO design because many of the analyzed tracts ($N = 195$) are not contiguous to other sampled tracts. It is thus not clear what a spatial model would accomplish when less than 25% of the total tracts in the city ($N = 865$) are studied, effectively censoring the majority of contiguous spatial units. Nevertheless, modeling spatial dependence between tracts might reduce "noise" introduced by spatially correlated errors. Information about spatial dependence might also make it possible to obtain reasonable measures of neighborhood disorder even for areas sparsely assessed by direct observation.

of individual-level properties are well established in the behavioral sciences, the development of ecometric research procedures and statistical methods for the study of ecological and other macrolevel units is in its infancy. We have explored the systematic social observation (SSO) of public spaces linked to neighborhood surveys, census data, and police records in an effort to create reliable and valid measures of neighborhood-level disorder. The SSO is an especially important case for "ecometrics" in sociology given the potential utility of videotaping as an observational strategy in the study of neighborhoods and other collectivities.

Consider some of the benefits for neighborhood research designs. Unlike on-site assessments, videotapes can be stored and made available for future researchers. They can be retrieved and revisited on demand, whether for assessing interrater reliability, for recoding, or to construct new variables not considered by the original investigators. An even richer possibility concerns the generative potential of video. Suppose that one finds an unusual concentration of criminal activity in certain neighborhoods that cannot be easily explained on the basis of measurable demographic characteristics or social processes like collective efficacy. Researchers can go back to the videos for selected face blocks to look anew at what might lie behind the density of criminal activity. Although not used as such in the present article, this possibility for video-based SSO is similar to the use of ethnography to generate hypotheses for future inquiry.[39]

Another advantage pertains to the multilevel assessment of neighborhood effects in sociological and developmental research. One of the central problems of extant research is that neighborhoods are treated as static constructs even though we know that neighborhoods change, often rapidly. Moreover, individuals move frequently during the course of longitudinal studies, yet typically, census tract data are "assigned" to individuals based on their past residence. Videotaping offers a relatively cheap and effective way to track within-neighborhood change and changes in the neighborhoods of residence. For example, as individuals move, interviewers at follow-up contacts could adopt an SSO strategy while in the process of carrying out individual-level assessments. Videotape records could then be matched to prior tapes and coded for change. When integrated with census data and perhaps surveys, the possibilities for discovering the ways in which neighborhoods influence individual development are greatly enhanced. Moreover, because face blocks are the first unit of measurement, "neighborhoods" can later be defined by the researcher at varying levels

[39] Beyond the scope of this article are potential ethical dilemmas in the use of systematic videotaping procedures, such as privacy and informed consent by street participants ("human subjects"?), the unique identification of individuals, and the appropriation of videotapes by the police for use as evidence of crimes.

of aggregation. Based on theory or emergent findings, for example, adjacent face blocks might be pieced together to form ecological units that better conform to the processes at hand. In this way, the use of SSO at the microlevel of blocks offers maximum flexibility for neighborhood-level research.

Our techniques for SSO also forge an explicit link with the technological advances that are transforming the ways in which research is conducted. Advertisers, fund-raisers, and market researchers are far ahead of sociologists in their command of geographic databases to cull sociological information and construct profiles of community process. With face blocks as the unit, SSO video can be linked to geographic information system (GIS) databases that allow for instantaneous merger with rich sources of information. For instance, SSO could in principle be linked to address-level databases on employment, real estate sales, and building code violations.

Finally, and perhaps most important, SSO provides the sights, sounds, and feel of the streets. Much as practiced by the original Chicago school of urban sociology (Abbott 1997), SSO takes researchers to the streets in a very real way. The present article has only scratched the surface of the potential for such a take on community. We thus believe the final story line pertains not just to collective efficacy or public disorder as theoretical constructs, but the potential scope of systematic social observation as an analytic tool. Visual cues are salient in many dimensions of social life; systematically observing them in their natural social context should be, as Reiss (1971) and Whyte (1988) argued, and a generation of students of the city before them (e.g., Park and Burgess 1921), a fundamental part of the sociological enterprise.

APPENDIX

A Model for Uncertainty in Systematic Social Observation

To estimate the magnitude of measurement errors in scales derived from SSO at the NC ($N = 80$) level, Raudenbush and Sampson (1999*b*) formulated a three-level model. We adapt that model to the data at hand, with multiple items nested within face blocks nested within census tracts. Level 1 captures the error due to item inconsistency within a face block; level 2 captures error due to variation between face blocks within census tracts while also adjusting for time-of-day effects; and level 3 specifies the "true-score" variation, that is, the variation between census tracts ($N = 196$) on the latent variable of interest. The model produces reliability estimates and the estimated correlation between latent variables adjusting for measurement error at each level. Because the item responses are dichotomous (presence or absence of each indicator of disorder within a face block), the model is a three-level logistic regression model.

Let Y_{ijk} take on a value of unity if indicator i of disorder is found present in face block j of tract k, with $Y_{ijk} = 0$ if not; and let μ_{ijk} denote the probability $Y_{ijk} = 1$. As is standard in logistic regression, we define η_{ijk} as the log-odds of this probability. Thus, we have

$$Y_{ijk}|\mu_{ijk} \sim \text{Bernoulli};$$

$$E(Y_{ijk}|\mu_{ijk}) = \mu_{ijk}$$

$$Var(Y_{ijk}|\mu_{ijk}) = \mu_{ijk}(1 - \mu_{ijk}).$$

$$\eta_{ijk} = \log\left(\frac{\mu_{ijk}}{1 - \mu_{ijk}}\right).$$

At level 1, the log-odds of finding disorder on item i depends on which aspect of disorder is of interest (physical or social) and which specific item is involved. Let D_{pijk} take on a value of 1 if item i is an indicator of physical disorder, 0 otherwise; and let $D_{sijk} = 1$ similarly take on a value of 1 if that item indicates social disorder. Then we have

$$\eta_{ijk} = D_{pijk}\left(\pi_{pjk} + \sum_{m=1}^{9} \alpha_{mjk} X_{mijk}\right) + D_{sijk}\left(\pi_{sjk} + \sum_{m=1}^{6} \delta_{mjk} Z_{mijk}\right),$$

where X_{mijk}, $m = 1, \ldots, 9$ are dummy variables representing nine of the ten items that measure physical disorder (each taking on a value of 1 or 0); Z_{mijk}, $m = 1, \ldots, 6$ are dummy variables representing six of the seven items that measure social disorder. We "center" each X and Z around its grand mean, which leads to the following definitions: π_{pjk} is the adjusted log-odds of finding physical disorder on a "typical item" when observing face block j of tract k; π_{sjk} is the adjusted log-odds of finding social disorder on a "typical item" when observing face block j of tract k; α_{mjk} reflects the "severity" level of item m within the physical disorder scale;[40] similarly, δ_{mjk} reflects the "severity" level of item m within the social disorder scale.[41] The item difficulties alpha and delta could, in principle, be allowed to vary across face blocks or NCs; however, in the absence of theory that

[40] Interpreting these coefficients in terms of "severity" requires that they be multiplied by -1 (Raudenbush and Sampson 1999*b*).

[41] One benefit of this model is that face-block measures of disorder, π_{pjk} and π_{sjk}, are adjusted for missing data. Because of the expense of coding the videotapes, we coded a random subsample of face blocks within tracts. Face blocks not sampled have data from the observation log but not the coding log. No bias arises because the coded face blocks constituted a representative sample of face blocks in the tracts. Nevertheless, controlling the item difficulties enables all of the data to be effectively used in the analysis (Raudenbush and Sampson 1999*b*).

might predict such variation, they will be held constant in the interest of parsimony. Thus, $\alpha_{mjk} = \alpha_m$ and $\delta_{mjk} = \delta_m$ for all j, k.

The level-2 model accounts for variation between face blocks within tracts on latent face-block disorder. Each is predicted by the overall tract-level disorder and the time of day during which the face block was observed:

$$\pi_{pjk} = \beta_{pk} + \sum_{q=1}^{5} \theta_{pqk}(\text{time})_{qjk} = u_{pjk},$$

$$\pi_{sjk} = \beta_{sk} + \sum_{q=1}^{5} \theta_{sqk}(\text{time})_{qjk} = u_{sjk},$$

$(\text{time})_{qjk}$ for $q = 1, \ldots, 5$ are five time-of-day indicators. They indicate the hours 7:00 to 8:59 A.M.; 9:00 to 10:59 A.M.; 11:00 A.M. to 12:59 P.M.; 1:00 to 2:59 P.M.; and 3:00 to 4:59 P.M., where the omitted group is from 5:00 to 6:59 P.M.; θ_{pqk} and θ_{sqk} are regression coefficients that capture the time-of-day effects on observing physical and social disorder within tract k. These could be allowed to vary over tracts, but for parsimony we hold them constant such that $\theta_{pqk} = \theta_{pq}$ and $\theta_{sqk} = \theta_{sq}$ for all k. Here β_{pk} and β_{sk} are the "true" scores for tract k on physical and social disorder, respectively, adjusting for time of day. The random effects u_{pjk}, u_{sjk} are assumed bivariate normally distributed with zero means, variances τ_{pp} and τ_{ss}, and covariance τ_{ps}. The variances will be large when face blocks vary greatly within tracts on disorder.

The third and final level of the model describes variation between tracts, the key units of neighborhood measurement, on physical and social disorder. We have simply

$$\beta_{pk} = \gamma_p + \upsilon_{pk},$$

$$\beta_{sk} = \gamma_s + \upsilon_{sk},$$

where γ_p and γ_s are the grand mean levels of physical and social disorder in Chicago neighborhoods and the random effects υ_{pk} and υ_{sk} are assumed bivariate normally distributed with zero means, variances ω_{pp} and ω_{ss}, and covariance ω_{ps}. The variances will be large when tracts vary greatly on their levels of disorder.

Measurement Results

The three models were combined as described in Raudenbush and Sampson (1999*b*), and all model parameters were estimated simultaneously by penalized quasi likelihood (Breslow and Clayton 1993) using an algorithm

TABLE A1

Level-1 HLM Results for Variation in SSO Disorder Items within Face Blocks: Item Difficulty

Item	Coefficient	SE
Physical disorder:		
Intercept	−2.30**	.10
Cigarettes, cigars in street/gutter	3.36**	.03
Garbage/litter on street/sidewalk	2.33**	.03
Empty beer bottles visible in street	1.11**	.03
Tagging graffiti	.33**	.04
Graffiti painted over	(reference item)	
Gang graffiti	−.68**	.05
Abandoned cars	−1.22**	.05
Condoms on sidewalk	−2.63**	.08
Needles and syringes	−2.90**	.09
Political message graffiti	−5.08**	.29
Social disorder:		
Intercept	−6.96**	.19
Adults loitering or congregating	3.83**	.24
People drinking alcohol	.61*	.29
Peer group, gang indicators present	(reference item)	
People intoxicated	−.06	.34
Adults fighting or hostilely arguing	−.41	.37
Prostitutes on street	−.49	.38
People selling drugs	−.81	.42

* $P < .05$.
** $P < .01$.

described in detail by Raudenbush (1995) and now implemented in version 4 of the HLM program (Bryk et al. 1996). Table A1 presents the estimates for the item response model. Items with negative coefficients have low probabilities of occurrence and thereby are rarer and more “severe” than are items with positive coefficients. In the physical disorder scale, the presence of cigarettes or cigars on the street or sidewalk, garbage and litter, along with the presence of empty beer bottles are comparatively less severe than the presence of gang graffiti, abandoned cars, or condoms on the sidewalk. Item severity thus conforms to intuitive expectations. The exception is political graffiti, which is exceptionally rare yet not generally regarded as especially severe. The item severities for physical disorder vary substantially, a feature of a well-behaved scale. In contrast, the item severities in the social disorder scale are clumped at the severe end except for adults loitering or congregating and drinking. This pattern reflects the low frequency of the social disorder indicators. Although the item severities are not well separated, their ordering does correspond to theoretical

TABLE A2

LEVEL-2 HLM RESULTS FOR VARIATION IN SSO DISORDER ITEMS BETWEEN FACE BLOCKS WITHIN TRACTS: TIME OF DAY EFFECTS

Time	Coefficient	SE
Physical disorder:		
7:00–8:59	.25**	.05
9:00–10:59	.09**	.03
11:00–12:59	.10*	.04
1:00–2:59	.15**	.04
3:00–4:59	.05	.04
5:00–6:59	(reference time)	
Social disorder:		
7:00–8:59	−.72**	.20
9:00–10:59	−.80**	.13
11:00–12:59	−.53**	.17
1:00–2:59	−.00	.14
3:00–4:59	−.09	.12
5:00–6:59	(reference time)	

* $P < .05$.
** $P < .01$.

expectation, with adults loitering and drinking alcohol being less severe than adults fighting, prostitution, or drug sales.

Table A2 provides estimates of the effects of time of day. Presumably social interactions in public view occur with relatively little frequency early in the morning and more frequently later on. This appears to be true of those social interactions involving some element of disorder as well. Note the mainly positive trend in time for social disorder, with coefficients of −.72, −.80, −.53, −.00, and −.09 as the day progresses. As expected, no such trend is apparent in the case of physical disorder. All model estimates in this article are adjusted for time-of-day effects.

Variance-Covariance Components and Reliability

The estimation of the variance-covariance components in table A3 provides information that we use to assess the quality of SSO measures. For comparative purposes, we report results for both the 196 census tracts and the larger NCs ($N = 80$). The intratract correlations are .46 and .98 for physical and social disorder, respectively. Formally,

$$\rho_{TRp} = \frac{\omega_{pp}}{\omega_{pp} + \tau_{pp}}; \; \rho_{TRs} = \frac{\omega_{ss}}{\omega_{ss} + \tau_{ss}}.$$

TABLE A3

Level-3 HLM Results: Variance-Covariance Components and Measurement Properties for Disorder Scales at Face Block and Tract Levels

	Physical Disorder		Social Disorder
Between face blocks:			
Variance	.67**		.02
	(.02)		(.16)
Covariance		.06	
		(.05)	
Reliability	.37		.00
Between census tracts:			
Variance	.56**		1.12**
	(.06)		(.16)
Covariance		.47**	
		(.08)	
Reliability	.95		.72

* $P < .05$.
** $P < .01$.

Thus, the intratract correlations, ρ_{TRp} and ρ_{TRs} for physical and social disorder, respectively, are the ratios of variance between tracts to overall variance, after adjusting for item inconsistency. The intra-NC correlation for physical disorder is .39 (Raudenbush and Sampson 1999*b,* pp. 28–29). The fact that the intratract correlation is larger than the intra-NC correlation arises because census tracts are internally more homogeneous than are NCs. The very high intratract correlation of .98 for social disorder suggests that there is little reliable variation between face blocks within tracts. Because evidence of social disorder is rare, data on social disorder at the face-block level are sparse, leading to imprecise estimation of the within-tract variance component.

Closely related to the intratract correlation is the internal consistency reliability of tract-level measurement. The latter depends on the intratract correlation but also on the number of face blocks sampled and the item severities. If all items are “severe,” evidence of disorder is hard to find, and reliability will be comparatively low. Raudenbush and Sampson (1999*b*) formulate the reliability for tract *k,* in the case of physical disorder, as given by

$$\lambda_{pk} = \frac{\omega_{pp}}{\omega_{pp} + \frac{\tau_{pp}}{J_k} + \frac{1}{nJ_k w}},$$

where λ_{pk} is the reliability of the physical disorder measure for tract *k;* $n = 10$ is the number of items in the scale; J_k = the number of face blocks sampled in tract *k;* and w = the average value of the estimates of $\mu_{ijk}(1 - \mu_{ijk})$ in tract *k*. The formula for social disorder is analogous, revealing that reliability will be high when (*a*) the between-tract variance ω_{pp} is large relative to the within-tract variance τ_{pp}; (*b*) the number of items in the scale, *n,* is large; (*c*) the number of face blocks sampled, J_k, is large; and (*d*) the probability of finding an item of disorder in a given face block (μ_{ijk}) is near .50, at which point w achieves its maximum.

The results in table A3 show that for physical disorder, the average of these reliabilities is .95. This extremely high reliability reflects a reasonably high intra–face-block correlation of .46, the large number of face blocks per tract, and the well-behaved item severities noted earlier. The average reliability for social disorder is lower than for physical disorder but still relatively high (.72). The difference reflects the extreme rarity of many of the indicators of social disorder and exists despite the high intratract correlation and the large number of face blocks per tract (98). Interestingly, the reliabilities for the 80 NCs studied by Raudenbush and Sampson (1999*b*) are not much different: .98 and .83 for physical disorder and social disorder, respectively. As tracts afford much greater power than NCs to detect between-neighborhood effects and are also more homogeneous internally, the similarities in reliabilities provide further evidence supporting the use of tracts as units of analysis. What about even lower levels of aggregation? Unfortunately, but not surprisingly given our results, reliabilities at the face-block level are simply unacceptable: .37 for physical disorder and .00 for social disorder.

REFERENCES

Abbott, Andrew. 1997. "Of Time and Space: The Contemporary Relevance of the Chicago School." *Social Forces* 75:1149–82.

Bellair, Paul. 1997. "Informal Surveillance and Street Crime: Disentangling Reciprocal Effects." Paper presented at the annual meeting of the American Sociological Association, Toronto, August.

Belsley, David, Edwin Kuh, and Roy Welsch. 1980. *Regression Diagnostics: Identifying Influential Data and Sources of Collinearity.* New York: Wiley.

Brantingham, Patricia, and Paul Brantingham. 1995. "Environment, Routine, and Situation: Toward a Pattern Theory of Crime." *Advances in Criminological Theory* 5:259–94.

Breslow, N., and D. G. Clayton. 1993. "Approximate Inference in Generalized Linear Mixed Models." *Journal of the American Statistical Association* 88:9–25.

Bryk, Anthony, and Stephen Raudenbush. 1992. *Hierarchical Linear Models: Applications and Data Analysis Methods.* Newbury Park, Calif.: Sage.

Bryk, Anthony, Stephen Raudenbush, Richard Congdon, and M. Seltzer. 1996. *HLM: Hierarchical Linear and Nonlinear Modeling with the HLM/2L and HLM/3L Programs.* Chicago: Scientific Software International.

Bursik, Robert J., Jr. 1988. "Social Disorganization and Theories of Crime and Delinquency: Problems and Prospects." *Criminology* 26:519–52.
Bursik, Robert J., Jr., and Harold Grasmick. 1993. *Neighborhoods and Crime: The Dimensions of Effective Community Control.* New York: Lexington.
Carr, Patrick J. 1998. "Keeping Up Appearances: Informal Social Control in a White Working-Class Neighborhood in Chicago." Ph.D. dissertation. University of Chicago, Department of Sociology.
Carter, Woody, Jody Dougherty, and Karen Grigorian. 1996. "Videotaping Neighborhoods." Working paper. University of Chicago, National Opinion Research Center.
Cohen, Lawrence, and Marcus Felson. 1979. "Social Change and Crime Rate Trends: A Routine Activity Approach." *American Sociological Review* 44:588–608.
Cohen, Lawrence, James Kluegel, and Kenneth Land. 1981. "Social Inequality and Predatory Criminal Victimization: An Exposition and Test of a Formal Theory." *American Sociological Review* 46:505–24.
Covington, Jeanette, and Ralph B. Taylor. 1991. "Fear of Crime in Urban Residential Neighborhoods: Implications of Between- and Within-Neighborhood Sources for Current Models." *Sociological Quarterly* 32:231–49.
Farrington, David. 1979. "Longitudinal Research on Crime and Delinquency." Pp. 289–348 in *Crime and Justice,* vol. 1. Edited by Norval Morris and Michael Tonry. Chicago: University of Chicago Press.
Felson, Marcus. 1987. "Routine Activities and Crime Prevention in the Developing Metropolis." *Criminology* 25:911–31.
Fisher, Franklin, and Daniel Nagin. 1978. "On the Feasibility of Identifying the Crime Function in a Simultaneous Model of Crime Rates and Sanction Levels." Pp. 361–99 in *Deterrence and Incapacitation: Estimating the Effects of Criminal Sanctions on Crime Rates,* edited by Alfred Blumstein, Jacqueline Cohen, and Daniel Nagin. Washington, D.C.: National Academy Press.
Garofalo, James, and John H. Laub. 1978. "The Fear of Crime: Broadening Our Perspective." *Victimology* 3:242–53.
Goffman, Erving. 1963. *Behavior in Public Places: Notes on the Social Organization of Gatherings.* New York: Free Press.
Gottfredson, Michael, and Travis Hirschi. 1990. *A General Theory of Crime.* Stanford, Calif.: Stanford University Press.
Greenberg, Stephanie, and W. Rohe. 1986. "Informal Social Control and Crime Prevention in Modern Urban Neighborhoods." Pp. 79–118 in *Urban Neighborhoods: Research and Policy,* edited by Ralph B. Taylor. New York: Praeger.
Hagan, John, and Bill McCarthy. 1997. *Mean Streets: Youth Crime and Homelessness.* New York: Cambridge University Press.
Hagan, John, and Ruth Peterson, eds. 1995. *Crime and Inequality.* Stanford, Calif.: Stanford University Press.
Hanushek, Eric, and J. Jackson. 1977. *Statistical Methods for Social Scientists.* New York: Academic.
Hearn, Frank. 1997. *Moral Order and Social Disorder: The American Search for Civil Society.* New York: Aldine de Gruyter.
Hunter, Albert. 1985. "Private, Parochial and Public Social Orders: The Problem of Crime and Incivility in Urban Communities." Pp. 230–42 in *The Challenge of Social Control: Citizenship and Institution Building in Modern Society,* edited by Gerald Suttles and Mayer Zald. Norwood, N.J.: Ablex Publishing.
Jacobs, Jane. 1961. *The Death and Life of Great American Cities.* New York: Random House.
Janowitz, Morris. 1975. "Sociological Theory and Social Control." *American Journal of Sociology* 81:82–108.
Jöreskog, Karl G., and Dag Sörbom. 1993. *LISREL 8: Structural Equation Modeling.* Chicago: Scientific Software International.

Kasarda, John, and Morris Janowitz. 1974. "Community Attachment in Mass Society." *American Sociological Review* 39:328–39.
Kelling, George. 1998. "Crime Control, the Police, and Culture Wars: Broken Windows and Cultural Pluralism." Paper presented at the National Institute of Justice, Washington, D.C., December.
Kelling, George, and Catherine Coles. 1996. *Fixing Broken Windows: Restoring Order and Reducing Crime in Our Communities.* New York: Free Press.
King, Gary, Robert Keohane, and Sidney Verba. 1993. *Designing Social Inquiry: Scientific Inference in Qualitative Research.* Princeton, N.J.: Princeton University Press.
Kornhauser, Ruth Rosner. 1978. *Social Sources of Delinquency: An Appraisal of Analytic Models.* Chicago: University of Chicago Press.
Land, Kenneth, Patricia McCall, and Lawrence Cohen. 1990. "Structural Covariates of Homicide Rates: Are There Any Invariances across Time and Space?" *American Journal of Sociology* 95:922–63.
Liska, Allen, and Paul Bellair. 1995. "Racial Composition and Crime: Convergence over Time." *American Journal of Sociology* 101:578–610.
Liska, Allen, and Barbara Warner. 1991. "Functions of Crime: A Paradoxical Process." *American Journal of Sociology* 96:1441–63.
Lofland, Lyn. 1973. *A World of Strangers: Order and Action in Urban Public Space.* New York: Basic Books.
Mazerolle, Lorraine Green, Colleen Kadleck, and Jan Roehl. 1998. "Controlling Drug and Disorder Problems: The Role of Place Managers." *Criminology* 36:371–404.
Morenoff, Jeffrey. 1999. "Unraveling Paradoxes in Public Health: Race and Ethnic Group Differences in Pregnancy Outcomes and the Neighborhood Context of Health." Ph.D. dissertation. University of Chicago, Department of Sociology.
Morenoff, Jeffrey, and Robert J. Sampson. 1997. "Violent Crime and the Spatial Dynamics of Neighborhood Transition: Chicago, 1970–1990." *Social Forces* 76:31–64.
NORC (National Opinion Research Center). 1995. PHDCN Project 4709. Systematic Social Observation Coding Manual, June 1995. Chicago: NORC.
Park, Robert E., and Ernest Burgess. 1921. *Introduction to the Science of Sociology.* Chicago: University of Chicago Press.
Pattillo, Mary E. 1998. "Sweet Mothers and Gangbangers: Managing Crime in a Black Middle-Class Neighborhood." *Social Forces* 76:747–74.
Perkins, Douglas, John Meeks, and Ralph B. Taylor. 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.
Perkins, Douglas, and Ralph B. Taylor. 1996. "Ecological Assessments of Community Disorder: Their Relationship to Fear of Crime and Theoretical Implications." *American Journal of Community Psychology* 24:63–107.
Perkins, Douglas, A. Wandersman, R. Rich, and Ralph B. Taylor. 1993. "The Physical Environment of Street Crime: Defensible Space, Territoriality and Incivilities." *Journal of Environmental Psychology* 13:29–49.
Raudenbush Stephen W. 1995. "Hierarchical Linear Models to Study the Effects of Social Context on Development." Pp. 165–201 in *The Analysis of Change,* edited by J. Gottman. Hillsdale, N.J.: Lawrence Erlbaum.
Raudenbush, Stephen, and Robert J. Sampson. 1999*a*. "Assessing Direct and Indirect Effects in Multilevel Designs with Latent Variables." *Sociological Methods and Research.* In press.
———. 1999*b*. " 'Ecometrics': Toward a Science of Assessing Ecological Settings, with Application to the Systematic Social Observation of Neighborhoods." *Sociological Methodology* 29:1–41.
Reiss, Albert J., Jr. 1971. "Systematic Observations of Natural Social Phenomena." Pp.

3–33 in *Sociological Methodology,* vol. 3, edited by Herbert Costner. San Francisco: Jossey-Bass.

Rountree, Pamela, and Kenneth Land. 1996. "Perceived Risk versus Fear of Crime: Empirical Evidence of Conceptually Distinct Reactions in Survey Data." *Social Forces* 74:1353–76.

Sampson, Robert J. 1988. "Local Friendship Ties and Community Attachment in Mass Society: A Multi-Level Systemic Model." *American Sociological Review* 53: 766–79.

Sampson, Robert J., and Dawn Jeglum-Bartusch. 1998. "Legal Cynicism and Subcultural (?) Tolerance of Deviance: The Neighborhood Context of Racial Differences." *Law and Society Review* 32:777–804.

Sampson, Robert J., Stephen Raudenbush, and Felton Earls. 1997. "Neighborhoods and Violent Crime: A Multilevel Study of Collective Efficacy." *Science* 277:918–24.

Sewell, William, Jr. 1992. "A Theory of Structure: Duality, Agency, and Transformation." *American Journal of Sociology* 98:1–29.

Shaw, Clifford R., and Henry D. McKay. 1942. *Juvenile Delinquency and Urban Areas.* Chicago: University of Chicago Press.

Short, James F., Jr., and Ivan Nye. 1957. "Reported Behavior as a Criterion of Deviant Behavior." *Social Problems* 5:207–13.

Singer, Simon. 1981. "Homogeneous Victim-Offender Populations: A Review and Some Implications." *Journal of Criminal Law and Criminology* 72:779–88.

Skogan, Wesley. 1990. *Disorder and Decline: Crime and the Spiral of Decay in American Cities.* Berkeley: University of California Press.

Skogan, Wesley, and Susan Hartnett. 1998. *Community Policing, Chicago Style.* New York: Oxford University Press.

Skogan, Wesley, and Michael Maxfield. 1981. *Coping with Crime: Individual and Neighborhood Reactions.* Beverly Hills, Calif.: Sage Publications.

Spergel, Irving. 1995. *The Youth Gang Problem: A Community Approach.* New York: Oxford University Press.

Taub, Richard, D. Garth Taylor, and Jan Dunham. 1984. *Paths of Neighborhood Change: Race and Crime in Urban America.* Chicago: University of Chicago Press.

Taylor, Ralph B. 1995. "Street Blocks with More Nonresidential Land Use Have More Physical Deterioration: Evidence from Baltimore and Philadelphia." *Urban Affairs Review* 31:20–36.

———. 1997. "Social Order and Disorder of Street Blocks and Neighborhoods: Ecology, Microecology, and the Systemic Model of Social Disorganization." *Journal of Research in Crime and Delinquency* 34:113–55.

———. 1999. "The Incivilities Thesis: Theory, Measurement, and Policy." In *Measuring What Matters.* Washington, D.C.: National Institute of Justice, in press.

Taylor, Ralph B., Stephen Gottfredson, and Sidney Brower, 1984. "Block Crime and Fear: Defensible Space, Local Social Ties, and Territorial Functioning." *Journal of Research in Crime and Delinquency* 21:303–31.

Taylor, Ralph, Sally Shumaker, and Stephen Gottfredson. 1985. "Neighborhood-Level Links between Physical Features and Local Sentiments: Deterioration, Fear of Crime, and Confidence." *Journal of Architectural Planning and Research* 21:261–75.

Tilly, Charles. 1973. "Do Communities Act?" *Sociological Inquiry* 43:209–40.

Whittemore, Alice S. 1989. "Errors-in-Variables Regression Using Stein Estimates." *American Statistician* 43:226–28.

Whyte, William F. 1943. *Street Corner Society: The Social Structure of an Italian Slum.* Chicago: University of Chicago Press.

Whyte, William H. 1988. *City: Rediscovering the Center.* New York: Doubleday.

Wikström, Per-Olof. 1991. *Urban Crime, Criminals, and Victims.* New York: Springer-Verlag.

Wilson, James Q., and George Kelling. 1982. "The Police and Neighborhood Safety: Broken Windows." *Atlantic* 127:29–38.

Wilson, William Julius. 1987. *The Truly Disadvantaged: The Inner City, the Underclass, and Public Policy.* Chicago: University of Chicago Press.

Wolfgang, Marvin. 1958. *Patterns in Criminal Violence.* Philadelphia: University of Pennsylvania.

Wright, Richard, and Scott Decker. 1997. *Armed Robbers in Action: Stickups and Street Culture.* Boston: Northeastern University Press.