Journal of Criminal Justice 38 (2010) 758–766

Contents lists available at ScienceDirect

Journal of Criminal Justice

Revisiting Broken Windows Theory: Examining the Sources of the Discriminant Validity of Perceived Disorder and Crime☆

Jacinta M. Gau [a,*], Travis C. Pratt [b]

[a] Department of Criminal Justice, California State University, San Bernardino, 5500 University Pkwy., San Bernardino, CA 92407, United States
[b] School of Criminology and Criminal Justice, Arizona State University, 411 N. Central Ave., Phoenix, AZ 85004-4420, United States

A B S T R A C T

Two lines of critiques have developed in reference to broken windows theory: (1) Concentrated disadvantage appears to be more intricately linked with disorder than the theory allows for; and (2) There is concern that disorder and crime lack discriminant validity in that people do not actually distinguish between the two. The present study integrated these two perspectives by examining whether concentrated disadvantage—including disorder itself—affects the extent to which people view disorder and crime as separate problems. Multivariate models showed that people who believe their neighborhood to be more disorderly were more likely to make distinctions between disorder and crime. Theoretical recommendations for future tests of broken windows theory are presented and the policy implications for order maintenance policing programs are discussed.

Introduction

There has been a recent upswing in the production of empirical tests of propositions specified by broken windows theory. The broken windows custom has traditionally assumed that disorder is a construct quite apart from crime and that disorder temporally precedes crime in a causal fashion. Failing to address disorderly conditions in certain areas has been hypothesized to spark a wave of serious crime within those needy neighborhoods or communities (Kelling & Coles, 1996; Skogan, 1990; Wilson & Kelling, 1982). This assumption, however, has been called into question by more recent empirical analyses (e.g., Sampson & Raudenbush, 1999). One emerging issue concerns the subjectivity of disorder as perceived by persons living in neighborhoods; specifically, the question has been raised as to whether people see disorder as a distinct problem separate from other negative neighborhood conditions (e.g., crime) or whether they view disorder as part-and-parcel to general area malaise. To this end, recent research has shown that disorder and crime have either marginal discriminant validity (Armstrong & Katz, 2010; Worrall, 2006a) or none at all (Gau & Pratt, 2008). There is, therefore, doubt as to just how different disorder and crime actually are.

Compounding the issue of discriminant validity (or the lack thereof) is the related debate about the origins of disorder and which coexisting signals of community disarray might influence citizens' perceptions of the amount of disorder in their neighborhoods. Sampson and Raudenbush (2004) concluded that area levels of socioeconomic disadvantage and racial heterogeneity influenced people's perceptions of disorder even more so than actual, observed levels of disorder did. This contradicted Wilson and Kelling's (1982) proposition that disorder is an exogenous construct that has independent effects on how people feel about their neighborhoods.

With this backdrop in mind, the present study addressed the discriminant validity of disorder and crime with a focus on the effect that people's perceptions of noxious neighborhood conditions, including disorder itself, have on these individuals' likelihood of perceiving disorder and crime as either separate or as the same. The results have implications for both criminological theory and crime control policy; specifically, evidence that the discriminant validity of disorder and crime is contingent upon local conditions—possibly contingent upon the prevalence of disorder itself—would suggest that incivilities and related problems are highly context-specific and should be addressed accordingly.

Broken Windows Theory: Disorder, Crime, and Discriminant Validity

Wilson and Kelling (1982) were not the first to point out the deleterious effects that disorder can have on communities, but they were the first to accuse disorder of actually causing crime. They hypothesized that even a single instance of disorder (the metaphorical "broken window") can spark a chain reaction of community decline if it is not fixed immediately (see also Skogan, 1990). This logic applied to everything from vandalism to obnoxious teenagers to

☆ This paper was accepted under the Editorship of Kent Joscelyn.
* Corresponding author.
E-mail address: jgau@csusb.edu (J.M. Gau).

0047-2352/$ – see front matter © 2010 Elsevier Ltd. All rights reserved.
doi:10.1016/j.jcrimjus.2010.05.002

pushy panhandlers. Wilson and Kelling believed that the failure to address these problems in a timely manner fostered a belief among community residents that all mechanisms of formal and informal social control had failed. Residents would eventually cede streets, parks, and other public spaces to the criminals who saw the lack of cohesiveness and control as a prime opportunity to practice their trades.

While this idea sounds good, empirical research on broken windows theory has been equivocal. Some studies (Savolainen, 2007; Skogan, 1990; Xu, Fiedler, & Flaming, 2005) have proffered support; others, however, allow only the tentative conclusion that some types of disorder may be related to some types of crimes in some areas using some—but not other—measurement techniques (Brown, Perkins, & Brown, 2004; Kurtz, Koons, & Taylor, 1998; Sampson & Raudenbush, 1999; Taylor, 2001; Wilcox, Quisenberry, Cabrera, & Jones, 2004). In addition, the precise relationship between the two is unclear even when a link is uncovered. Disorder and crime do often co-occur, but it is difficult to say whether this overlap comports with broken windows' sequential, causal, disorder-to-crime process or, alternatively, whether disorder and crime are both sub-components of larger conditions of concentrated sociostructural disadvantage.

The validity of broken windows theory matters greatly from a policy perspective because broken windows theory has had dramatic impact on the field of policing. In short, Wilson and Kelling (1982) put police in charge of cleaning up the streets. Police, according to broken windows, could keep serious crime at bay by maintaining control over disorderly conditions. Various pieces of, and spin-offs from, broken windows theory have been adopted by police departments nationwide (Harcourt, 2001; Harcourt & Ludwig, 2006; Kelling & Coles, 1996). Evaluations of order maintenance policing strategies have produced mixed findings. Some have shown support for these strategies (Braga et al., 1999; Corman & Mocan, 2005; Kelling & Sousa, 2001; Sampson & Cohen, 1988; Worrall, 2006b; see also Skogan, 2008; also see Cerda et al., 2009, finding a possible small effect, but noting its sensitivity to specification error) and others have not (Harcourt & Ludwig, 2006; Hinkle & Weisburd, 2008; Katz, Webb, & Schaefer, 2001; Novak, Hartman, Holsinger, & Turner, 1999; see also Eck & Maguire, 2000; Gau & Brunson, 2010; Greene, 1999). The prevalence and popularity of broken windows-type policing interventions necessitate further testing of the theory upon which these policies are premised because it is unclear at this point whether or not the theory has merit and, if it does, under what conditions order maintenance might be a viable approach.

Within the mix of challenges to and critiques of broken windows theory, two separate but potentially-related themes have emerged. The first theme centers on the discriminant validity (i.e., the empirical separability) of disorder and crime. The broken windows thesis makes several implicit assumptions about the nature of disorder and crime. One pivotal assumption is that disorder and crime are actually different constructs; that is, that there is a clear line differentiating each one from the other. This assumption is critical to the theory because disorder cannot cause crime if disorder *is* crime; such a state of affairs would render the theory fatally tautological. Broken windows' assumption of separability is logically problematic from the outset because many of the actions that the theory characterizes as types of disorder—such as prostitution, public drinking, and vandalism—are actually low-level forms of crime (Gau & Pratt, 2008; Sampson, 2006; Sampson & Raudenbush, 1999). Thus, there is an endogeneity problem built right into the broken windows framework.

Broken windows theory could withstand this criticism, however, if, irrespective of the philosophical debate over logical fallacies, citizens were to in fact make a mental distinction between disorder and crime. To this end, some studies have employed survey-based perceptual measures to test for discriminant validity between the disorder and crime constructs using confirmatory factor analyses designed to shed light on the underlying factor structure. Ross and Mirowsky (1999) assessed the discriminant and convergent validity of physical disorder, social disorder, and crime. Social and physical disorder did demonstrate convergent validity, which seemed consistent with broken windows, but the crime indicators loaded strongly on the disorder factor. This implies that the two constructs are not distinct, as broken windows theory maintains, but are, rather, two pieces of a larger whole. Worrall (2006a) and Armstrong and Katz (2010) found mixed results for the discriminant validity of perceptual incivilities, personal victimization, and perceived crime. Perceived crime and disorder appeared to load on a single factor, though there was an apparent distinction between physical incivilities and personal victimization. Overall, crime and disorder did not demonstrate clear or consistent discriminant validity. Gau and Pratt (2008) compared fit indices and the overall quality of one- and two-factor models to determine which model fit the data better. Although the model depicting disorder and crime as distinct constructs fit the data well, the unacceptably-high between-factor correlation ($r = .92$) led the authors to conclude that it is inappropriate and artificial to draw a line between disorder and crime.

A second issue that has been brought up in the literature concerns the effect that concentrated disadvantage can have on both crime and disorder. Broken windows theory manifests a singular focus on disorder and downplays the effect of other neighborhood problems. The decontextualization of disorder is a mark against the theory because negative environmental conditions are preeminent considerations in the analysis of communities and crime. The study of structural disadvantage has a long history in criminological research, dating back to Shaw and McKay's (1942) formulation of social disorganization theory, of which the relatively-recent broken windows theory is an offshoot (Pratt & Gau, 2009). Within the tenets of social disorganization theory (and its revised systemic version), concentrated social and economic disadvantage are strong—albeit indirect—determinants of area crime rates (Kornhauser, 1978; Lowenkamp, Cullen, & Pratt, 2003; Pratt & Cullen, 2005; Sampson & Groves, 1989). Disadvantage and crime, moreover, are not spread evenly across municipalities but instead tend to be concentrated in particular areas (Sherman, 1995; Sherman, Gartin, & Buerger, 1989), which results in differential exposure to these noxious conditions across the residents of a given city or town.

Research has shown that citizens' perceptions of disorder are affected by concentrated disadvantage. Sampson and Raudenbush (2004; see also Wilcox et al., 2004) found that neighborhood levels of poverty and racial heterogeneity in Chicago shaped residents' perceptions of the severity of disorder more so than disorder itself did. Franzini, Caughy, Nettles, and O'Campo (2007) found similar results in Boston, though poverty stood out in this analysis as the driving force behind disorder perceptions. Other researchers have linked disorder to area land use. Business-oriented land use increases both perceived (Wilcox et al., 2004) and independently-assessed incivilities (Kurtz et al., 1998; Sampson & Raudenbush, 1999). Overall, then, it appears that certain sociostructural conditions can shape the extent to which citizens believe disorder to be prevalent and/or problematic in their neighborhoods and communities (see also Piquero, 1999).

An ongoing debate that bears on the present analysis is that concerning the "best" method for measuring disorder (Hipp, 2007; Skogan, 1990). Some researchers (Sampson, 2001; Sampson & Raudenbush, 1999) have criticized what they believe to be an overreliance on perceptual measures and have advocated a switch to independent observation by trained research personnel as a more objective—and hence more accurate—approach to assessing the concentration of disorder in an area. While there are clear merits to diversifying the approach to measuring disorder, perceptual measures remain integral to the study of broken windows theory because this theory is premised upon people's perceptions of, and reactions to,

disorder in their neighborhoods and communities. Independently-observed measures quantify amounts and types of disorder, but they fail to tap the psychological and interpretive impact that disorder has on the people who must navigate incivilities in their daily lives. In broken windows research, then, as in any area of criminological research, the specific question at issue in a study should guide the selection of the appropriate disorder measure. Given the present study's focus on citizens' perceptions of disorder, perceptual measures were opted for because independently-observed disorder would not have permitted a satisfactory test of the hypothesis under investigation.

Current Focus

Two general statements sum up the extant research on broken windows theory. First, there is a relationship between disadvantage and people's perceptions of disorder. People's psychological constructions of disorder do not flow linearly or unambiguously from actual neighborhood disorder; other area conditions profoundly influence people's estimates of the quantity and severity of disorder (Franzini et al., 2007; Sampson & Raudenbush, 2004). Second, it is unclear whether disorder and crime are truly different phenomena or whether they are, instead, the same thing in the minds of a community's residents (Armstrong & Katz, 2010; Gau & Pratt, 2008; Ross & Mirowsky, 1999; Worrall, 2006a). What is missing from the literature is a bridge between these two areas of study to address whether people's beliefs about the state of their neighborhood affect the discriminant validity of indicators of disorder and crime. Specifically, the presence and extent of disadvantage in a neighborhood could influence residents' assessments of disorder and crime and could determine whether they view disorder as a problem apart from crime or whether they, conversely, view disorder and crime both as components of the overarching negative conditions in their area of residence.

The results of this study have implications for the applicability of broken windows theory in neighborhoods of varying levels of (dis) advantage. Wilson and Kelling (1982) made much use of the language of prevention in their original proposal—they wrote of the need for police and private citizens to keep disorder at bay and to not allow it to infiltrate the city. They also described the possibility that in some places, disorder and crime may already be so prevalent that controlling disorder is no longer an option because serious crime has already befallen the area. This implies that broken windows theory (and its policing strategy) may be most applicable in relatively disorder-free neighborhoods where disorder can still be held in check. It remains to be seen, however, how people's perceptions of the state of their neighborhoods affect the distinction they make between evidence of disorder and instances of actual crime.

Methods

Data

Data for this study were drawn from a large district comprising the eastern portion of Washington State. The variables used here were a subset of items from a larger survey designed to assess respondents' perceptions of crime and disorder, their fear of crime, their attitudes toward local police, and their beliefs about social cohesion in their neighborhoods and communities of residence. The survey was funded with federal grant money and was a joint project between the U.S. attorney's office for this district and the major university in the area. The area is a mix of mid-sized cities and small towns. The population is primarily White. This largely nonurban environment provided an opportune setting for this research because most prior studies of broken windows have focused on urban areas (e.g., Sampson & Raudenbush, 1999, 2004; Skogan, 1990; Taylor, 2001) to the near exclusion of nonurban jurisdictions (but see Gau & Pratt, 2008). This has led to questions concerning the applicability and validity of broken windows theory outside of inner-cities and other densely-populated areas (e.g., Crank, Giacomazzi, & Heck, 2003; Reisig & Cancino, 2004).

The survey methodology was a panel design consisting of surveys conducted in 2003 and 2006 and followed Dillman's (2000) total design method for mail surveys. Surveys and postage-paid return envelopes were mailed to a random sample of addresses in each of the twenty-one largest municipalities in each of the counties. A second wave was sent to nonresponders from all counties, and targeted third and fourth waves were sent to counties that maintained response rates below 20 percent even after multiple mailings. The original sample consisted of 2,879 returned, usable surveys, for a 32.5 percent response rate. This response rate is fairly typical of the current response rates researchers are achieving in mail and telephone surveys (Curtin, Presser, & Singer, 2005; Gau & Pratt, 2008; Holtfreter, Reisig, & Pratt, 2008; McCarty, House, Harman, & Richards, 2006). Researchers in other criminological fields of study have been forced to settle for even lower response rates (see, e.g., Tyler & Wakslak, 2004, with a 22 percent response rate) but have provided important contributions nonetheless. A prior study using the 2003 data set probed for nonresponse bias and found no evidence of a problem (Gau & Pratt, 2008).

A question on the first survey asked respondents if they would be willing to participate in another survey of the same sort, and those who marked "yes" became the sampling frame for the second survey in 2006. The second survey achieved a 76.6 percent response rate. The data set for the second survey was merged with the first according to identification numbers linked to mailing addresses, and demographic information was cross-checked to ensure that the particular respondents who filled out the first survey were the same people who filled out the second survey. There were 1,029 cases in the final sample of usable panel data.

The panel design of this survey made the data well-suited for a broken windows-based test concerning disorder and crime. Broken windows theory is fundamentally a longitudinal theory, as it posits events and consequences that unfold over time. Most tests of the theory, however, have used cross-sectional data (e.g., Sampson & Raudenbush, 1999, 2004; Xu et al., 2005; see Brown et al., 2004; Taylor, 2001 for exceptions), which has limited some of the conclusions that could be drawn from these analyses because disorder and crime tend to co-occur, making it difficult to establish temporal ordering when the constructs are measured simultaneously. In the present study, the dependent variable was taken from the second survey and all independent variables were drawn from the first survey. Temporal ordering is, therefore, inherent in the survey design.

Dependent Variable

The dependent variable was a *disorder-crime difference score* and came from the second survey. Respondents were provided a list of crime and disorder problems and were asked to rate each one on a one-to-four scale where one indicated "no problem" and four represented "a serious problem." The survey instrument directed respondents to limit their assessments to their neighborhoods of residence, which comported with prior researchers' contentions that disorder should be measured at small units of analyses like neighborhoods (e.g., Perkins & Taylor, 1996). The time respondents spend outside their neighborhoods is inestimable, whereas relative confidence can be placed in their knowledge about the area immediately surrounding their places of residence. It is important to note that despite the fact that respondents were randomly selected for participation on the basis of their residence within the cities or towns of interest, it is their perceptions of their neighborhoods—not their broader communities—that are under examination here. The

survey methodology did not permit respondents to be linked to neighborhoods, but the item wording implicitly embedded a neighborhood-level consideration within the data. The analysis was therefore able to avoid making the untenable assumption that perceptions of disadvantage are constant across any given city or town.

The first step was to construct disorder and crime scales from the survey items. Items were placed in the scales on the basis of prior literature in order to maintain consistency with existing broken windows studies, though it is worth noting that prior studies have revealed little agreement regarding the mix of items that should be used to construct disorder scales. Face validity and prior research were relied upon to construct the present disorder scale. Some of the survey items (such as garbage/litter, stray animals, and groups of loiterers) were clearly indicators of disorder and were automatically included in the scale. Other items (gangs, prostitution, drug use, and drug sales) were more ambiguous, so prior research was consulted to make a decision about whether to consider them disorder or crime. Several prior studies concerning disorder and crime within a broken windows framework have used items tapping illegal drug use, or physical evidence thereof, as disorder measures (Doran & Lees, 2005; Gau & Pratt, 2008; McGarrell, Giacomazzi, & Thurman, 1997; Ross & Mirowsky, 1999; Sampson & Raudenbush, 1999, 2004; Skogan, 1990; Worrall, 2006a). Gangs have also typically been defined as disorder (Gau & Pratt, 2008; McGarrell et al., 1997; Piquero, 1999; Reisig & Parks, 2004; Sampson & Raudenbush, 1999; Skogan, 1990). Finally, Worrall (2006b) included prostitution in his disorder scale. Based on precedent, then, the gang, drug, and prostitution items were included in the disorder scale. The scales demonstrated internal consistency (Cronbach's $\alpha_{disorder} = .886$; $\alpha_{crime} = .879$) and unidimensionality (see Appendix A for factor loadings).

Next, the scales were standardized to account for the fact that the disorder scale contained more items than the crime scale did. The two scales correlated at .799 ($p < .01$), which was expected given the doubts about discriminant validity. This strong correlation indicated the presence of substantial covariation between the scales: where disorder was high, so too was crime, yet the relationship was certainly not perfect. In the last step, the dependent variable, disorder-crime difference, was created by subtracting the standardized crime scale from the standardized disorder scale and then taking the absolute values of the difference scores.[1] Higher scores on the dependent variable represented a larger difference between disorder and crime because the scale increased as the gap between crime and disorder widened. The scale was reasonably normally distributed, so no transformations were made.

Independent Variables

Several independent variables were included that tapped different facets of concentrated disadvantage. Some of these were survey-based and others were derived from official sources. The two survey-based, individual-level independent variables were *disorder* and *crime*. These scales were the first survey's analogues to the second survey's disorder and crime scales that were used to construct the dependent variable. Perceived crime and disorder are measures of the extent to which people see their neighborhoods as troubled. As detailed above, it is important to include perceptual measures of neighborhood problems because, ultimately, it is people's feelings about these problems that influence those individuals' attitudes and behaviors. Both of these variables were mean-centered to correct for collinearity due to a high interitem correlation ($r = .752$, $p < .001$; items and factor loadings located in Appendix A).

Municipality-level census and uniform crime report variables were used to account for actual disadvantage in communities and to offer an alternative to the perceptual items described above so that both perceived and official measures were included in this study. Disorder, as a form of structural disadvantage, may be more likely to occur in areas beset by other social, structural, and economic problems. This would affect the distribution of perceived disorder across neighborhoods and communities. Because there was no way to parse respondents by neighborhood, the city/town level was the smallest unit of analysis available for the concentrated disadvantage variables. The variables were *low socioeconomic status* and *disadvantage*. They were composite scores created on the basis of content similarity, correlations, and factor loadings. *Low socioeconomic status* was a combination of the percent in each city that was below poverty and the percent that was unemployed. *Disadvantage* combined city crime rates (gathered from the target state's uniform crime report repository and computed as all index crimes per 1,000 citizens), the percent of each city's population aged 25 and over that did not possess a high school diploma, and the percent of the city's housing units that were vacant. All variables in both scales were standardized and then summed. Four respondent-level demographic variables were also entered as controls. These items were *age* (in years), *race* (0 = White; 1 = non-White), *sex* (0 = male; 1 = female), and *income* (1 to 10 scale).

Analytic Strategy

The analysis proceeded in two stages. First, an ordinary least squares (OLS) regression model[2] was estimated with the full sample to test for the effects of concentrated disadvantage and individual-level perceptions of disorder and crime on the disorder-crime difference score. Second, the sample was divided roughly in half according the disorder scale's median, and separate OLS regression models were estimated for each subsample to test for an interaction effect between disorder and the discriminant validity of disorder and crime. Confidence intervals were computed to determine whether there was a statistically significant difference between the disorder slopes under each of the two disorder conditions. Missing data were treated with simple listwise deletion because the maximum amount of data missing across the independent and dependent variables was 8 percent. When so little data is missing, listwise deletion is an acceptable strategy (Byrne & Watkins, 2003).

Results

Respondents' demographic features are located in Table 1. Approximately 32 percent of the respondents were female and the mean age was just under fifty-eight years (median age for the state = 36.8). Just under six percent of respondents were Nonwhite and while a greater representation of minorities would have been preferable, the White-dominated sample was actually not dramatically different from the census-based estimates of minority composition in the area where the survey was conducted. The city-level minority concentration ranged from 2.80 percent to 47.20 percent, with a mean of 18.15 (sd = 9.45) percent. In six percent of the cities, minority citizens comprised less than 10 percent of the total population.

Table 1
Descriptive Statistics for Variables used in the Analysis

Item or Scale	Minimum – Maximum	Mean (sd)
Disorder-Crime Difference Scale (DV)	.01 – 20.32	3.73 (3.06)
Disorder Scale (IV)	10 – 38	15.79 (5.77)
Crime Scale (IV)	5 – 19	7.08 (2.46)
Low SES	-2.08 – 3.68	.00 (1.57)
Disadvantage	-4.01 – 6.17	.00 (2.14)
Age	19 – 96	57.84 (14.11)
Nonwhite	0 – 1	.06 (na)
Female	0 – 1	.32 (na)
Income	1 – 10	md = 6.00

These differences between the sample and population demographics were primarily due to nonresponse decisions made by individuals, rather than to a flaw in the survey methodology which, as aforementioned, used random sampling and multiple waves to boost responses as much as possible. To assuage any concerns of potential nonresponse bias, though, correlations were estimated between respondents' demographics and the crime, disorder, and difference scales. The correlation matrix is located in Appendix B. The correlations between demographic characteristics and the key variables were small. Many achieved statistical significance, but this is not unexpected with a large sample. Income correlated with disorder and crime somewhat, which is predictable because people with higher incomes probably live in more socioeconomically-advantaged neighborhoods (see Sampson & Raudenbush, 1999; Wilcox et al., 2004; see also McGarrell et al., 1997). Age was also related somewhat to disorder and to the difference score; however, the correlation was weak and the two variables shared only 1.3 percent of their variance. The correlations that did surface function as reminders that the results of this analysis should be interpreted with an understanding of their potential limitations.

All models were checked for multicollinearity and autocorrelation. Collinearity diagnostics indicated no problem once disorder and crime were both centered on their respective means. All variance inflation factors were below 2.5 and the highest condition index was 17.7. It was necessary to check for possible autocorrelation because the individual survey respondents were nested within the twenty-one cities in the analysis. Durbin-Watson statistics for all models were approximately 1.9, indicating the absence of harmful autocorrelation.

Table 2 contains the results for the full-sample OLS model. The model was decent but not particularly robust, with 16.9 percent of the variation in the disorder-crime difference score explained. The disadvantage scales ("low SES" and "disadvantage") were nonsignificant, but this could have been the result of fact that these two variables were measured at the city level. City-level data can mask important between-neighborhood differences in socioeconomic status and crime rates. Future research should continue work on disaggregating these measures and using smaller units of analysis.

The only significant independent variable was perceived disorder. This measure emerged as an important predictor of the discriminant validity of disorder and crime ($\beta = .331$). The more disorder people saw in their neighborhoods, the better they differentiated between disorder and crime. It would appear from this finding that people's day-to-day exposure to disorder strongly affects their overall perceptions of both disorder and crime.

The strong effect of disorder on the discriminant validity of disorder and crime warranted further examination, so the sample was divided along the median of the first survey's disorder scale and the regression model was re-estimated independently for each subsample. This procedure functioned as a test for interaction effects to determine whether residing in an area that one considers to be of "low" or "high" disorder significantly affected the relationship between disorder and the difference score. Table 3 displays the results for the split-sample regression models.

Model 1 in Table 3 contains the results for the OLS regression model based on those respondents whose disorder scale scores fell below the scale median (median = 14.000). Model 2, by comparison, was estimated using those respondents whose disorder scores fell above the scale median. The 95 percent confidence intervals are also reported to facilitate comparisons between the two sets of coefficients. The "low disorder" model was not good. The R^2 showed that this set of independent variables explained only 2.1 percent of the variation in the difference score and the nonsignificant F-value meant the model was not useful. The coefficient for perceived disorder, nonetheless, was statistically significant and substantively meaningful ($\beta = -.117$). The negative sign of this coefficient indicated that when people did not consider their neighborhoods to be very disorderly, increases in that initially-low level of disorder caused a reduction in respondents' ability to tell disorder from crime.

The "high disorder" model again confirmed the importance of disorder in the discriminant validity of disorder and crime; however, this model differed importantly from Model 1 in that the predictors explained 26 percent of the variation in the disorder-crime difference score. Disorder again led with the strongest coefficient ($\beta = .440$), which was much larger in this model than in Model 1 ($\beta = -.117$). The sign, moreover, was positive, whereas the sign for the low disorder model was negative. Where people's baseline levels of disorder were high, increases in disorder were associated with better disorder-crime discriminant validity. This confirmed the presence of an interaction between baseline levels of disorder and respondents' ability to

Table 2
Full-Sample OLS Regression Model Predicting the Disorder-Crime Difference Score

Independent Variable	b (SE)	β
Disorder	.175* (.026)	.331
Crime	.080 (.061)	.065
Low SES	.014 (.063)	.007
Disadvantage	.028 (.046)	.020
Age	-.011 (.008)	-.048
Nonwhite	.014 (.439)	.001
Female	.219 (.216)	.033
Income	-.061 (.040)	-.053
Constant	4.673* (.578)	
N = 825		
R^2 = .169		
F = 20.786*		

* p<.001.

Table 3
OLS Regression Models for the Sample Split into Low and High Disorder

Independent Variable	*Model 1: Low Disorder* b (SE)	β	95% CI	*Model 2: High Disorder* b (SE)	β	95% CI
Disorder	-.148* (.064)	-.117	(-.274, -.022)	.338** (.045)	.440	(.249, .427)
Crime	.047 (.082)	.029	(-.114, .207)	.042 (.086)	.029	(-.126, .210)
Low SES	.057 (.052)	.052	(-.046, .159)	-.070 (.127)	-.025	(-.319, .108)
Disadvantage	.010 (.040)	.012	(-.069, .090)	.082 (.088)	.042	(-.091, .256)
Age	.001 (.007)	.006	(-.013, .015)	-.030* (.014)	-.104	(-.057, -.003)
Non-White	-.382 (.422)	-.043	(-1.212, .447)	.060 (.756)	.004	(-1.426, 1.547)
Female	.217 (.199)	.053	(-.174, .608)	.155 (.382)	.019	(-.596, .907)
Income	-.005 (.036)	-.007	(-.077, .066)	-.143* (.071)	-.094	(-.283, -.003)
Constant	2.555** (.575)		(1.424, 3.686)	5.116** (1.006)		(3.183, 7.093)
	N = 443			N = 382		
	R^2 = .021			R^2 = .258		
	F = 1.151 (ns)			F = 16.184**		
	mean disorder = 11.596			mean disorder = 20.540		
	mean crime = 5.723			mean crime = 8.592		

*p<.05 ** p<.001.

discriminate between disorder and crime. Where people's base levels of perceived disorder were low, an increase in perceived disorder actually led to a blending of disorder and crime, but where the base levels were high, increases resulted in more separation. These findings are discussed in detail below.

Discussion

Broken windows theory has caused a fair amount of controversy. Supporters tout the studies (e.g., Kelling & Coles, 1996; Kelling & Sousa, 2001; Skogan, 1990) that seem to illustrate the theory's validity, while critics point to the studies (e.g., Gau & Pratt, 2008; Harcourt, 2001; Sampson & Raudenbush, 1999; Taylor, 2001) that seem to demonstrate the opposite. Prior studies have linked area levels of structural and social disadvantage to disorder and crime in a manner suggesting that disorder may simply be one component of a larger condition of disadvantage. At the same time, questions have been raised as to whether disorder and crime may actually be the same thing, with disorder being but a less serious form of its criminal analogue. The purpose of the present study was to integrate these two perspectives. Perceived disorder and area levels of sociostructural disadvantage and were hypothesized to predict people's ability to separate disorder from crime; in other words, it was expected that the general state of an area would affect people's judgments about disorder and crime. Accordingly, given the results presented here, two conclusions deserve elaboration.

First, Wilson and Kelling (1982) hinted at a nonlinear trend in disorder when they suggested the presence of a "tipping point" (p. 38) at which disorder becomes so bad that order maintenance policing is no longer a viable strategy for crime reduction. The present study uncovered nonlinearity in a manner somewhat different from, but nonetheless consistent with, that predicted by Wilson and Kelling. In particular, both main and interaction effects were uncovered between perceived disorder and the disorder-crime difference score. In the full-sample model (Table 2), perceived disorder bore a strong and significant relationship to the disorder-crime difference scale, indicating that as people saw more disorder in their neighborhood, their ability to discern more routine signs of disorder from actual instances of crime improved. People who did not consider disorder to be very problematic in their neighborhoods, on the other hand, tended to collapse the two conditions into a single one. Dividing the sample into respondents who perceived "low" and "high" levels of disorder in their neighborhoods produced evidence of an interaction between disorder and the disorder-crime difference score. People who characterized the amount of disorder in their neighborhoods as minimal saw less difference between disorder and crime as disorder increased. The effect was exactly the opposite for those residing in comparative disorder: people with higher baseline levels of perceived disorder experienced increases in discriminatory capacity with every incremental increase in disorder. In short, people in orderly areas could not tell the difference, but people in disorderly areas could.

The most significant implication for broken windows theory and research is that there is systematic patterning to the discriminant validity of disorder and crime. Prior researchers who found little or no discriminant validity (Gau & Pratt, 2008; Ross & Mirowsky, 1999; Worrall, 2006a) did not investigate the potential for separability to vary interspatially. The present analyses indicate that there is variation in the degree to which citizens see disorder and crime as separate social problems. This variation, moreover, is not random—it is systematically tied to the contextual conditions people find themselves in. When citizens are accustomed to cleanliness and order, an increase in disorder could trigger fear and a sense that something in the neighborhood has gone terribly wrong. In these places, disorder may truly be a sign to residents that community control is slipping.

In disorderly areas, by contrast, residents may develop a sharper grasp on the distinction between routine annoyance and true danger because their constant exposure to disorder inoculates them against the fear that formal and informal controls are failing and that crime is on the rise. An increase in disorder may be viewed simply an increase in disorder, accompanied with few or no inferences about crime. These results support prior researchers' discoveries of ecological patterning in disorder types and quantities (e.g., Spelman, 2004) and underscore the importance of studying disorder at the smallest unit of analysis available (see also Perkins & Taylor, 1996).

Second, the results present a mixed bag of implications for broken windows-style policing strategies. The original tipping point idea is supported to the extent that when baseline levels of perceived disorder are low, order maintenance policing might effectively control citizens' fear that crime is on the rise. In low-disorder areas, shooing away aggressive panhandlers may have the same effect on residents' fear as apprehending serious criminal offenders. People who do not think their neighborhoods are disorderly may not make a conscious, qualitative distinction between the panhandler and the serious criminal. The present study supports the broken windows proposition that orderly areas should be kept orderly. Order maintenance policing may have promise in these areas.

The situation is very different, however, for people who see their neighborhoods as disorderly. In the present study, these people appeared to make conscious distinctions between disorder and crime; in fact, the discriminant validity of disorder and crime increased as perceived disorder rose. It may be that the more someone sees disorder on a routine basis, the better that person is equipped to distinguish between disorderly conditions and criminal acts. In short, people who consider their neighborhoods disorderly *do* see a difference between the panhandler and the serious criminal. Disorder may be overwhelming and intimidating to those who are not accustomed to it, but those who are familiar with it seem to have an easier time recognizing which actions or conditions are threatening to their safety and which are not.

High-disorder areas present a dilemma for police because it is not clear whether their efforts are best concentrated on disorder or on crime. Wilson and Kelling (1982) thought that disorderly areas would be too overrun with crime for police to realistically allot resources to disorder-reduction. Even if police did have the resources to fight disorder and crime simultaneously, order-maintenance strategies might fail because, in a manner of speaking, the cat's out of the bag—the disorder-to-crime sequence has already transpired and it is too late for prevention. Our results emphasize the additional need for police to consider citizens' perceptions of crime and disorder in their neighborhoods and their beliefs about which problem should be the primary target of police efforts. Anti-disorder campaigns might make residents of less disorderly areas think that the police are effective at controlling both disorder and serious crime, but it might make people in more disorderly areas wonder why the police are pouring their efforts into disorder when crime is such a problem. Residents of disorderly areas undoubtedly wish that something would be done about the noxious conditions of their neighborhoods (see Skogan, 2008), but it does not necessarily follow that they want the police to be the ones to do the clean-up. Citizens may see it as a job for public works and sanitation agencies rather than for law enforcement personnel. Even citizens who value community-based strategies sometimes still see traditional law enforcement as the key police function (Webb & Katz, 1997). In essence, no policing strategy works all the time in every location; solutions must be tailored to fit local problems (Sherman, 1986; Weisburd & Eck, 2004).

Proactive policing structured under the broken windows rubric has also been associated with negative consequences for police and communities in precisely those areas where it is generally employed most vigorously: disorderly, low-income, often minority neighborhoods and communities. Here, police are more likely to rely on

proactive strategies like widespread stops and frisks (Brunson & Miller, 2006; Weitzer, 2000). Minority citizens tend to bear the brunt of proactive policing, as evidenced by higher rates of stops and frisks relative to Whites (Fagan & Davies, 2000). Proactive order maintenance can create friction between police and citizens, the latter of whom feel that their rights and dignity are threatened when officers persistently scrutinize them (Gau & Brunson, 2010), which can undermine police-community relationships (Reisig, Bratton, & Gertz, 2007; Reisig & Parks, 2000, 2004; Reisig & Stroshine Chandek, 2001). This is particularly troublesome because citizens of lower-income neighborhoods generally have little or no recourse for expressing their grievances when they feel that the police have wronged them (Brunson, 2007; Kane, 2002). Other collateral consequences include the impact that increases in arrests for minor disorder offenses have on local jail populations and court docket capacity. Greene (1999) compared New York City's famous order maintenance policing effort during the early 1990s to an alternative community-based strategy that was used in San Diego during the same time period. She found that although misdemeanor arrest rates rose sharply in New York City and actually declined in San Diego, the two cities experienced similar reductions in crime. As a whole, these studies show that order maintenance policing can actually be detrimental in disorderly, disadvantaged communities, and the present analysis supports and extends the contention that this style of policing should be employed in these areas with caution.

Limitations to the present study include the rural and semi-urban nature of the sample, the low response rate on the first survey, and the lack of "objective" baseline levels of disorder and crime. We see each of these limitations, particularly the third, as opportunities for future research and underscore the fact that the present study is the beginning, not the end, of the debate over this topic. Regarding the first limitation, future researchers are encouraged to conduct similar analyses with samples gathered from other areas of the country with varying levels of urbanicity/rurality. Concerns over the low response rate and potential bias resulting therefrom were assuaged by correlations showing miniscule relationships between respondents' demographic features and their responses on the key disorder and crime variables. Perhaps one of the largest limitations was that the respondents were overwhelmingly White, which was a function of both the mostly-White population of the survey location and of minority nonresponse. Future research should therefore consider over-sampling strategies to perhaps ensure greater representations of minorities. Finally, studies should endeavor to include baseline levels of disorder and crime. Both perceptual and independently-observed disorder measures have their weaknesses, but combined, they could offer strong tests of broken windows-related questions. The discussion about the relationship between actual and perceived disorder—and the effect of the former on the latter—is far from over.

In the end, both academics and policymakers should be open to a more nuanced conceptualization of disorder, crime, and citizen perceptions of each. These relationships are complex and are clearly linked to overarching sociostructural conditions (see also Sampson & Raudenbush, 1999). The present study's results imply that the viability of broken windows theory and policing is not uniform across different types of neighborhoods. Police resources are strained and police agencies cannot afford to squander money and time. It is worth considering from the outset whether there is an identifiable reason to believe a given policing strategy—order maintenance or otherwise—will work in their jurisdiction before that strategy is launched (see also Sherman, 1992, 2002; Weisburd & Eck, 2004). It is time to reconsider how broken windows-based policing can be used strategically.

Acknowledgement

The data for this project were gathered using funds provided by the Project Safe Neighborhoods federal grant number 5710-0040.

Appendix A. Items and Factor Loadings for the Disorder and Crime Scales

Scale and Items	Item Factor Loadings
Disorder Scale for Dependent Variable	$\alpha = .886$
People drinking or drunk in public	.693
Groups of teenagers or others hanging out and harassing people	.724
Stray or uncontrolled dogs and other animals running around	.533
People using illegal drugs	.810
Vandalism	.753
Traffic problems	.579
Garbage/Litter	.629
Gangs	.783
Drunk drivers	.697
Prostitution	.620
People buying and/or selling drugs	.804
Crime Scale for Dependent Variable	$\alpha = .879$
Violent crime	.885
Rape	.873
Gun violence	.872
Burglary	.749
Robbery	.804
Disorder Scale for Independent Variable	$\alpha = .871$
Stray dogs running around	.535
Drunk drivers on the road	.702
People drinking to excess in public	.717
Groups of teenagers hanging out and harassing people	.768
Youth gangs	.752
People using illegal drugs	.758
Vandalism	.740
Noise	.698
Traffic problems	.596
Garbage/litter	.641
Crime Scale for Independent Variable	$\alpha = .852$
People's homes being broken into and things being stolen	.692
People being robbed or having their purses/wallets taken	.695
Domestic/intimate partner violence	.708
Rape/sex crimes	.794
Child abuse	.720
Violent crime	.798
Gun violence	.795

Appendix B. Correlations for Nonresponse Bias

	1	2	3	4	5	6	7
1. Difference Score (DV)	1.000						
2. Disorder	.411**	1.000					
3. Crime	.329**	.752**	1.000				
4. Age	-.091**	-.112**	-.066*	1.000			
5. Female	.077**	.065*	.086**	-.068*	1.000		
6. Non-White	.045	.060	.019	-.092**	.021	1.000	
7. Income	-.130**	-.215**	-.170**	-.270**	-.159**	-.100**	1.000

*p<.05 **p<.01.

Notes

1. When left in its original units, the disorder-crime difference scale contained the conceptual and statistical problem that each of the endpoints or anchors represented a difference, with the middle of the scale representing no difference. This led to a situation in which the gap between disorder and crime first closed and then widened. Taking the absolute value of the scale remedied this. The scale starts at zero and increases as the absolute difference between disorder and crime becomes larger.

2. The nested nature of the data might seem to warrant a multilevel approach to modeling the hypothesized relationships; however, OLS was used instead for three reasons. The first reason was that there were only twenty-one cities in the sample,

which would effectively reduce the sample size to twenty-one in a multilevel modeling context. This would result in low degrees of freedom and, accordingly, a threat of low statistical power. The second reason was that the Durbin-Watson test for autocorrelation did not indicate the presence of correlated error terms that would artificially deflate standard errors and result in spuriously significant findings. Multilevel modeling is considerably more complex both statistically and substantively, so absent a clear need for this more complicated analysis, the more straightforward method is justified. Finally, the most appropriate Level 2 unit of analysis from a conceptual standpoint would be neighborhoods, but the data did not permit respondents to be partitioned into neighborhoods. Cities, then, would have been the Level 2 units, which would raise the aforementioned power problem and also run contrary to the findings presented here that show that city-level disadvantage does not predict the disorder-crime difference score. In addition, community-level variables can obscure important cross-neighborhood variation, so using cities/towns as the macrolevel units would have raised questions about the appropriateness of this choice.

References

Armstrong, T., & Katz, C. (2010). Further evidence on the discriminant validity of perceptual incivilities measures. *Justice Quarterly*, *27*(2), 280—304.

Braga, A. A., Weisburd, D. L., Waring, E. J., Mazerolle, L. G., Spelman, W., & Gajewski, F. (1999). Problem-oriented policing in violent crime places: A randomized controlled experiment. *Criminology*, *37*(3), 541—580.

Brown, B. B., Perkins, D. D., & Brown, G. (2004). Crime, new housing, and housing incivilities in a first-ring suburb: Multilevel relationships across time. *Housing Policy Debate*, *15*(2), 301—345.

Brunson, R. K. (2007). "Police don't like black people": African-American young men's accumulated police experiences. *Criminology & Public Policy*, *6*(1), 71—102.

Brunson, R. K., & Miller, J. (2006). Young black men and urban policing in the United States. *British Journal of Criminology*, *46*, 613—640.

Byrne, B. M., & Watkins, D. (2003). The issue of measurement invariance revisited. *Journal of Cross-Cultural Psychology*, *34*, 155—175.

Cerda, M., Tracy, M., Messner, S. F., Vlahov, D., Tardiff, K., & Galea, S. (2009). Misdemeanor policing, physical disorder, and gun-related homicide. *Epidemiology*, *20*(4), 533—541.

Corman, H., & Mocan, N. (2005). Carrots, sticks, and broken windows. *Journal of Law and Economics*, *48*, 235—266.

Crank, J. P., Giacomazzi, A., & Heck, C. (2003). Fear of crime in a nonurban setting. *Journal of Criminal Justice*, *31*, 249—263.

Curtin, R., Presser, S., & Singer, E. (2005). Changes in telephone survey nonresponse over the past quarter century. *Public Opinion Quarterly*, *69*, 87—98.

Dillman, D. A. (2000). *Mail and internet surveys: The tailored design method* (2nd ed.). New York: John Wiley Co.

Doran, B. J., & Lees, B. G. (2005). Investigating the spatiotemporal links between disorder, crime, and the fear of crime. *The Professional Geographer*, *57*(1), 1—12.

Eck, J. E., & Maguire, E. R. (2000). Have changes in policing reduced violent crime? An assessment of the evidence. In A. Blumstein & J. Wallman (Eds.), *The crime drop in America* (pp. 207—265). Cambridge, MA: Cambridge University Press.

Fagan, J. & Davies, G. (2000). Street cops and broken windows: Terry, race, and disorder in New York City. *28 Fordham Urban Law Journal 457*.

Franzini, L., Caughy, M. O., Nettles, S. M., & O'Campo, P. (2007). Perceptions of disorder: contributions of neighborhood characteristics to subjective perceptions of disorder. *Journal of Environmental Psychology*, *28*, 83—93.

Gau, J. M., & Brunson, R. K. (2010). Procedural justice and order maintenance policing: A study of inner-city young men's perceptions of police legitimacy. *Justice Quarterly*, *27*(2), 255—279.

Gau, J. M., & Pratt, T. C. (2008). Broken windows or window dressing? Citizens' (in)ability to tell the difference between disorder and crime. *Criminology & Public Policy*, *7*(2), 163—194.

Greene, J. A. (1999). Zero tolerance: A case study of police policies and practices in New York City. *Crime and Delinquency*, *45*(2), 171—187.

Harcourt, B. E. (2001). *Illusion of order: The false promise of broken windows policing.* Cambridge: Harvard University Press.

Harcourt, B. E., & Ludwig, J. (2006). Broken windows: New evidence from New York City and a five-city social experiment. *The University of Chicago Law Review*, *73*(1), 271—320.

Hinkle, J. C., & Weisburd, D. (2008). The irony of broken windows policing: A micro-place study of the relationship between disorder, focused police crackdowns and fear of crime. *Journal of Criminal Justice*, *36*, 503—512.

Hipp, J. R. (2007). Block, tract, and levels of aggregation: Neighborhood structure and crime and disorder as a case in point. *American Sociological Review*, *72*, 659—680.

Holtfreter, K., Reisig, M. D., & Pratt, T. C. (2008). Low self-control, routine activities, and fraud victimization. *Criminology*, *46*, 189—220.

Kane, R. J. (2002). The social ecology of police misconduct. *Criminology*, *40*(4), 867—896.

Katz, C. M., Webb, V. J., & Schaefer, D. R. (2001). An assessment of the impact of quality-of-life policing on crime and disorder. *Justice Quarterly*, *18*(4), 825—876.

Kelling, G. L., & Coles, C. M. (1996). *Fixing broken windows.* New York: Simon & Schuster.

Kelling, G. L., & Sousa, W. H., Jr. (2001). *Do police matter? An analysis of the impact of New York City's police reforms.* New York, NY: The Manhattan Institute.

Kornhauser, R. R. (1978). *Social sources of delinquency.* Chicago, IL: University of Chicago Press.

Kurtz, E. M., Koons, B. A., & Taylor, R. B. (1998). Land use, physical deterioration, resident-based control, and calls for service on urban streetblocks. *Justice Quarterly*, *15*(1), 121—149.

Lowenkamp, C. T., Cullen, F. T., & Pratt, T. C. (2003). Replicating Sampson and Groves's test of social disorganization theory: Revisiting a criminological classic. *Journal of Research in Crime and Delinquency*, *40*(4), 351—373.

McCarty, C., House, M., Harman, J., & Richards, S. (2006). Effort in phone survey response rates: The effects of vendor and client-controlled factors. *Field Methods*, *18*, 172—188.

McGarrell, E. F., Giacomazzi, A. L., & Thurman, Q. C. (1997). Neighborhood disorder, integration, and the fear of crime. *Justice Quarterly*, *14*(3), 479—497.

Novak, K. J., Hartman, J. L., Holsinger, A. M., & Turner, M. G. (1999). The effects of aggressive policing of disorder on serious crime. *Policing: An International Journal of Police Strategies & Management*, *22*(2), 171—190.

Perkins, D. D., & Taylor, R. B. (1996). Ecological assessments of community disorder: Their relationship to fear of crime and theoretical implications. *American Journal of Community Psychology*, *24*(1), 63—107.

Piquero, A. (1999). The validity of incivility measures in public housing. *Justice Quarterly*, *16*(4), 793—818.

Pratt, T. C., & Cullen, F. T. (2005). Assessing macro-level predictors and theories of crime: A meta-analysis. In M. Tonry (Ed.), *Crime and justice: Vol. 32* (pp. 373—450). Chicago, IL: University of Chicago Press.

Pratt, T. C., & Gau, J. M. (2009). Social disorganization theory. In J. H. Copes & V. Topalli (Eds.), *Criminology theory: The reader* (pp. 104—112). Boston, MA: McGraw-Hill.

Reisig, M. D., Bratton, J., & Gertz, M. G. (2007). The construct validity and refinement of process-based policing measures. *Criminal Justice & Behavior*, *34*(7), 1005—1028.

Reisig, M. D., & Cancino, J. M. (2004). Incivilities in nonmetropolitan communities: The effects of structural constraints, social conditions, and crime. *Journal of Criminal Justice*, *32*, 15—29.

Reisig, M. D., & Parks, R. B. (2000). Experience, quality of life, and neighborhood context: A hierarchical analysis of satisfaction with police. *Justice Quarterly*, *17*(3), 607—630.

Reisig, M. D., & Parks, R. B. (2004). Can community policing help the truly disadvantaged? *Crime & Delinquency*, *50*(2), 139—167.

Reisig, M. D., & Stroshine Chandek, M. (2001). The effects of expectancy disconfirmation on outcome satisfaction in police-citizen encounters. *Policing: An International Journal of Police Strategies & Management*, *24*(1), 88—99.

Ross, C. E., & Mirowsky, J. (1999). Disorder and decay: The concept and measurement of perceived neighborhood disorder. *Urban Affairs Review*, *34*(3), 412—432.

Sampson, R. J. (2001). Book review: Breaking away from broken windows: Baltimore neighborhoods and the nationwide fight against crime, grime, fear, and decline. *American Journal of Sociology*, *107*, 501—502.

Sampson, R. J. (2006). Collective efficacy theory: Lessons learned and directions for future inquiry. In F. T. Cullen, J. P. Wright, & K. R. Blevins (Eds.), *Taking stock: The status of criminological theory, advances in criminological theory: Vol. 15* (pp. 149—168). New Brunswick, NJ: Transaction.

Sampson, R. J., & Cohen, J. (1988). Deterrent effects of the police on crime: A replication and theoretical extension. *Law & Society Review*, *22*(1), 163—189.

Sampson, R. J., & Groves, W. B. (1989). Community structure and crime: Testing social-disorganization theory. *American Journal of Sociology*, *94*(4), 774—802.

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

Sampson, R. J., & Raudenbush, S. W. (2004). Seeing disorder: Neighborhood stigma and the social construction of "broken windows." *Social Psychology Quarterly*, *67*(4), 319—342.

Savolainen, J. (2007). Public disorder and business victimization: Findings from a survey of female entrepreneurs. *Crime Prevention and Community Safety*, *9*, 1—20.

Shaw, C. R., & McKay, H. D. (1942). *Juvenile delinquency in urban areas.* Chicago, IL: The University of Chicago Press.

Sherman, L. W. (1986). Policing communities: What works? In A. J. Reiss, Jr. & M. Tonry (Eds.), *Crime and justice: Communities and crime* (pp. 343—386). Chicago, IL: University of Chicago Press.

Sherman, L. W. (1992). Attacking crime: Police and crime control. In M. Tonry (Ed.), *Crime and justice: A review of the research* (pp. 159—230). Chicago, IL: The University of Chicago Press.

Sherman, L. W. (1995). Hot spots of crime and criminal careers of places. In J. E. Eck & D. Weisburd (Eds.), *Crime and place* (pp. 35—52). Monsey, NY: Criminal Justice Press.

Sherman, L. W. (2002). Evidence-based policing: Social organization of information for social control. In E. Waring & D. Weisburd (Eds.), *Advances in criminological theory: Crime and social organization* (pp. 217—248). New Brunswick, NJ: Transaction Books.

Sherman, L. W., Gartin, P. R., & Buerger, M. E. (1989). Hot spots of predatory crime: routine activities and the criminology of place. *Criminology*, *27*(1), 27—55.

Skogan, W. G. (1990). *Disorder and decline.* Berkeley: University of California Press.

Skogan, W. G. (2008). Broken windows: Why—and how—we should take them seriously. *Criminology & Public Policy*, *7*(2), 195—202.

Spelman, W. (2004). Optimal targeting of incivility-reduction strategies. *Journal of Quantitative Criminology*, *20*(1), 63—88.

Taylor, R. B. (2001). *Breaking away from broken windows.* Boulder, CO: Westview Press.

Tyler, T. R., & Wakslak, C. J. (2004). Profiling and police legitimacy: Procedural justice, attributions of motive, and acceptance of police authority. *Criminology*, *42*(2), 253—281.

Webb, V. J., & Katz, C. M. (1997). Citizen ratings of the importance of community policing activities. *Policing: An International Journal of Police Strategies and Management*, *20*(1), 7—23.

Weisburd, D., & Eck, J. E. (2004). What can police do to reduce crime, disorder, and fear? *Annals of the American Academy of Political and Social Science*, *593*, 42—65.

Weitzer, R. (2000). Racialized policing: Residents' perceptions in three neighborhoods. *Law & Society Review*, *34*(1), 129—155.

Wilcox, P., Quisenberry, N., Cabrera, D. T., & Jones, S. (2004). Busy places and broken windows?: Toward defining the role of physical structure and process in community crime models. *Sociological Quarterly*, *45*(2), 185−207.

Wilson, J. Q. & Kelling, G. L. (1982). The police and neighborhood safety: Broken windows. *The Atlantic Monthly, March,* 29-38.

Worrall, J. L. (2006a). The discriminant validity of perceptual incivility measures. *Justice Quarterly*, *23*(3), 360−383.

Worrall, J. L. (2006b). Does targeting minor offenses reduce serious crime? A provisional, affirmative answer based on an analysis of county-level data. *Police Quarterly*, *9*(1), 47−72.

Xu, Y., Fiedler, M. L., & Flaming, K. H. (2005). Discovering the impact of community policing: The broken windows thesis, collective efficacy, and citizens' judgment. *Journal of Research in Crime and Delinquency*, *42*(2), 147−186.