

Annual Review of Criminology

Looking Through Broken Windows: The Impact of Neighborhood Disorder on Aggression and Fear of Crime Is an Artifact of Research Design

Daniel T. O'Brien,^{1,2} Chelsea Farrell,³
and Brandon C. Welsh^{3,4}

¹School of Public Policy and Urban Affairs, Northeastern University, Boston, Massachusetts 02120, USA; email: d.obrien@neu.edu

²Boston Area Research Initiative, Northeastern and Harvard Universities, Boston, Massachusetts 02120, USA

³School of Criminology and Criminal Justice, Northeastern University, Boston, Massachusetts 02120, USA

⁴Netherlands Institute for the Study of Crime and Law Enforcement, 1091 Amsterdam, The Netherlands

Annu. Rev. Criminol. 2019. 2:53–71

First published as a Review in Advance on
September 26, 2018

The *Annual Review of Criminology* is online at
criminol.annualreviews.org

<https://doi.org/10.1146/annurev-criminol-011518-024638>

Copyright © 2019 by Annual Reviews.
All rights reserved

**ANNUAL
REVIEWS CONNECT**

www.annualreviews.org

- Download figures
- Navigate cited references
- Keyword search
- Explore related articles
- Share via email or social media

Keywords

broken windows theory, physical disorder, social disorder, neighborhood perceptions, aggression, fear of crime, neighborhood attachment, meta-analysis

Abstract

Broken windows theory (BWT) has heavily influenced social science and policy over the past 30 years. It posits that disorder in neighborhoods leads to elevated crime by inviting additional criminal activity and by discouraging the positive social behavior that prevents crime. Scholars have debated the veracity of BWT, and here we conduct a meta-analysis of 96 studies to examine the effects of disorder on residents' (*a*) general proclivities for aggressive behavior and (*b*) perceptions of and attitudes toward their neighborhood (e.g., fear of crime), with particular attention to aspects of research design that might confound causal inference. We found no consistent evidence that

disorder induces greater aggression or more negative attitudes toward the neighborhood. Studies that found such effects disproportionately utilized weaker research designs that omit key correlates or confound perceptions of disorder with other neighborhood attitudes. We explore implications for theory, research, and policy.

INTRODUCTION

Broken windows theory (BWT) has been one of the most influential concepts in the social and behavioral sciences over the past thirty years as well as one of the most controversial. First introduced by Kelling & Wilson (1982), BWT posits that the presence of physical and social disorder (e.g., graffiti, public drunkenness, panhandling, dilapidation) leads to the proliferation of more serious crime. This has inspired a prominent approach to law enforcement (i.e., broken windows policing) that emphasizes the enforcement of low-level misdemeanors as an effective way to diminish crime (Bratton & Knobler 2001, Kelling & Coles 1997), as well as numerous theoretical offshoots in disciplines beyond criminology, including public health (Cohen et al. 2000) and child development (Milam et al. 2010). This prominence has been accompanied by a great deal of controversy. Evaluations of the impact of broken windows policing on crime reduction have been inconsistent (Braga & Bond 2008, Braga et al. 2015, Harcourt & Ludwig 2006, Taylor et al. 2011, Weisburd et al. 2011). There is also evidence that the strategy might implicitly encourage racial profiling (Harcourt 2001, Taylor 2001). More to the point, there are questions of whether the basic premise of BWT even holds (Markowitz et al. 2001, O'Brien & Sampson 2015, Perkins et al. 1993, Sampson & Raudenbush 1999, Steenbeek & Hipp 2011, Taylor 2001), leaving an unclear picture of a popular perspective.

Testing whether disorder indeed leads to crime might be conducted at either of two levels. Much of the rhetoric to date has emphasized the ecological level, focusing either on disorder's relationship with neighborhood-level patterns of crime or whether broken windows policing does in fact diminish crime. But at its heart, BWT is a proposition about individual-level psychology, i.e., that disorder induces behaviors and attitudes that might contribute to increases in disorder and crime. Direct tests of this individual-level proposition are essential if we are to understand not only whether disorder leads to crime but also the mechanisms that drive such a relationship. This, in turn, sheds light on the kinds of policing strategies that might be most effective. There is a large literature that has tested this very question, although the methodologies and results have been varied. In an effort to make sense of this body of work, the current review conducts a meta-analysis that evaluates the overall evidence for the propositions of BWT at the individual level. As we do so, we closely examine (a) the different behavioral pathways by which disorder might lead to elevated crime and (b) the extent to which research design might influence the inferred magnitude of such effects.

DISORDER AND BEHAVIOR

Disorder has long been of interest to urban scholars, going back to the original studies of neighborhood variation in nineteenth-century London (Booth 1903, Mayhew 1862). It was also treated as an important indicator of neighborhood social quality by the Chicago School (Park et al. 1925/1984) and Jane Jacobs (1961). Kelling & Wilson's (1982) seminal essay diverged from previous scholars in that it was the first to attribute a causal role to disorder. They stated that the accumulation of disorder in a neighborhood directly leads to greater frequency and severity of uncivil and illegal

behavior, including violent crime, thereby instigating what Skogan (1986, 1990) has referred to as a cycle of disorder and decline. When framed this way, BWT is a model of ecological succession in which one particular set of conditions (i.e., disorder) determines the future trajectory of the neighborhood (i.e., increases in crime). This is a natural way to formulate BWT, as it is a theory of cross-neighborhood variations in crime, but it is somewhat incomplete. The ecological process at the heart of BWT—disorder leading to violent crime—must be mediated by the perceptions and behaviors of residents of and visitors to a neighborhood (Skogan 1992). Put another way, BWT depends on the psychological proposition that disorder induces attitudinal and behavioral responses that subsequently contribute to crime.

Kelling & Wilson (1982) were explicit about the psychological components of BWT, proposing two pathways by which disorder might elicit behaviors that result in more crime. The first, which described a direct effect between disorder and criminal behavior, was encapsulated in the allegory from which their theory takes its name: a broken window or other form of unchecked disorder is a signal that uncivil or delinquent acts will go unpunished, thereby encouraging such behavior. In so many words, broken windows beget more broken windows, and disorder begets additional violations of local norms and laws. Although this proposition is intuitively appealing, it was justified at the time by evidence from a single study with a rather weak quasi-experimental design. Zimbardo (1969) placed two seemingly abandoned vehicles in two quite different contexts: the Bronx, New York, and Palo Alto, California. Eventually, both cars were stripped for valuable parts and thoroughly demolished, indicating that otherwise upstanding citizens will become vandals if they believe that no one cares.¹

The second proposed pathway from disorder to crime was indirect, operating through those who are not inclined toward delinquent behavior. Disorder signals an unmanaged space where danger might lurk, leading residents to retreat from public life. This would erode the community's natural ability to manage behavior, therefore permitting crime to proliferate. This indirect pathway has received less attention over the years, likely because it is less evocative, but it actually was the one with greater empirical basis at the time of the essay's publication. Numerous studies had connected disorder to fear of crime (Garofalo & Laub 1978, Hunter 1974, Wilson 1975), and, possibly for this reason, Kelling & Wilson (1982, p. 31) spend at least as much time describing how disorder makes residents "think that crime, especially violent crime, is on the rise . . . They will use the streets less often, and when on the streets will stay apart from their fellows."

Experimental studies by psychologists have repeatedly offered evidence for the two behavioral pathways posited to underlie BWT. In a now-classic study, Cialdini et al. (1990) found that individuals were more likely to litter in spaces where litter had already been strewn than in clean ones. Similarly, a series of studies by Keizer et al. (2008, 2013) showed that individuals are likely to commit a variety of low-level transgressions when disorder is present. An important advance of this latter set of work in terms of BWT is the demonstration of cross-norm disinhibition, i.e., disorder of one type does not only elicit more of the same behavior but also encourages individuals to commit any number of transgressions that might present themselves. This is crucial to the premise that disorder in general leads to other, more serious delinquent acts. Regarding the second pathway, work by O'Brien and colleagues (O'Brien et al. 2014b, O'Brien & Wilson 2011) on community perception has shown that disorder acts as a salient heuristic when people evaluate the social quality of an unfamiliar neighborhood. Importantly, these inferences influence

¹To be fair, the process of destruction was not quite as automatic in the Palo Alto case, where the car sat untouched for a week until Zimbardo himself hit it with a sledgehammer. Within hours, the community had joined in and the car was flipped over and destroyed. Kelling & Wilson (1982) took this variation as further support of their theory; as long as the signal that no one was concerned about this piece of property is made sufficiently salient, people will participate in vandalism.

subsequent behavior, leading to lower levels of cooperation when in the presence of higher disorder. Subsequent studies using similar methodologies have replicated this finding (Yang & Pao 2015) and have even shown that the behavioral results can be generalized to disorder that is nonsocial (e.g., nonstraight edges, asymmetry) (Kotabe et al. 2016).

Although the experimental evidence for psychological responses to disorder is suggestive, it does not guarantee that those mechanisms are sufficiently strong to drive the neighborhood-level cycle of disorder and decline. First, many of the studies exposed naïve participants to unfamiliar neighborhoods, thereby forcing them to make inferences and decisions based on limited information. Resident judgments of their own neighborhood are likely to be rooted at least in part on a deeper base of knowledge and experience, relying less heavily on disorder (see O'Brien & Kauffman 2013). Second, the outcome measures of the experiments have been proxies of social behavior (e.g., behavioral economic games) and rather low-level incivilities, like littering, trespassing, and keeping small amounts of unattended money. For obvious ethical reasons, they have not been able to directly assess the tendency of disorder to encourage violent crime, which is one of the most basic aspects of BWT. These two weaknesses leave a gap in knowledge that can be filled only by studies that examine the effect of habitual exposure to disorder on residents' behaviors and attitudes. There are many such studies and they vary widely in their methodology, the outcomes they examine, and the level of support (or counterevidence) they provide for BWT. Here, we leverage meta-analysis to organize and make sense of this wide-ranging literature, evaluating the *in vivo* impact of disorder on behaviors that might lead to elevated crime and disorder in urban neighborhoods.

Current Study

The current review draws from a comprehensive database of studies that have tested the effects of disorder in a neighborhood on individual-level, resident outcomes. For our purposes, we limit our attention to those studies within the database that tested either of the two main pathways by which disorder is proposed to lead to crime. Namely, does disorder predict (*a*) increased proclivity for offending or aggressive behavior or (*b*) perceptions of and attitudes toward one's neighborhood, including fear of crime and neighborhood attachment. The primary goal of the meta-analysis is to assess the global effect of disorder for each of these pathways. Additionally, it allows us to account for variation in key study characteristics (i.e., research design) that might moderate the inferred relationship between disorder and a given outcome (Lipsey & Wilson 2001). This is critically important because, as others have noted (e.g., Kubrin 2008), there has been inconsistency over the years regarding the definition of disorder and the techniques that best measure it. To this latter point, we examine three methodological debates that have emerged and examine the extent to which each might influence results:²

1. Measurement of disorder: A study might measure disorder as social elements (e.g., drunkenness), physical elements (e.g., graffiti), or both. Researchers must also choose whether to measure it objectively through audits or perceptually through resident surveys. In theory, each of these different operationalizations might reflect a different form or experience of disorder that has distinct consequences for behaviors and attitudes.

²Some readers will note a fourth area of contention that we do not mention: the subjective nature of disorder, meaning its definition might shift across individuals and communities (see O'Brien et al. 2014b, O'Brien & Wilson 2011, Taylor 2005, Yang & Pao 2015). That said, a thorough analysis of how different groups perceive the elements of disorder would require its own extensive meta-analysis and cannot be easily incorporated into the current analysis.

2. Confounding perceived disorder with outcomes: Perceived disorder can be measured either by aggregating the responses of many residents or, more simply, as the perceptions of the same focal individual whose outcomes are being analyzed. This design, however, has an inherent weakness, specifically for tests of the second pathway. It might find evidence for a causal relationship between the experience of disorder and other attitudes about a neighborhood when in actuality both are manifestations of a single psychological construct for evaluating neighborhood conditions.
3. Inclusion of appropriate covariates: Disorder is highly correlated with numerous variables associated with both crime and fear of crime among residents. The most apparent of these is wealth. Another comes from one of the most trenchant critiques of BWT. Sampson & Raudenbush (1999) have provided empirical evidence that disorder does not cause crime but that both are symptoms of collective efficacy, or the ability of a community to draw from established relationships to enforce shared norms. We thus examine the extent to which the inclusion or omission in a statistical model of measures of wealth, collective efficacy, or related social processes (e.g., social cohesion, networks) influences the inferred effect of disorder on behavior.

METHODS

This section describes the methods used in this review, including the literature searches and meta-analytic procedures.

Literature Searches

The analysis utilizes a database of studies that test the effects of neighborhood disorder on individual-level outcomes of residents. We conducted searches of five leading electronic bibliographic databases [Web of Science (ISI), Criminal Justice Abstracts, Sociological Abstracts, PsycINFO, and Medline] for any of the following terms occurring in conjunction with the word “neighborhood” in the title, abstract, or text: “broken windows,” “social disorder,” “physical disorder,” “incivilities,” “physical deterioration,” and “structural deterioration.” To ensure thoroughness, we repeated this search by adding the terms “aggression” or “fear of crime” and removing the term “neighborhood.” The database was fully updated as of September 2016. Reviewing the results of these searches, we limited inclusion to those studies that (*a*) tested the effect of at least one operationalization of neighborhood disorder on (*b*) an individual-level outcome. We also used citations in these studies to identify any additional studies not uncovered in our searches. We categorized individual-level outcomes into nine groups: socio-behavioral outcomes, cognitive function, health-related behaviors, mental health, negative experiences, neighborhood perceptions, physical health, relations with the police, and miscellaneous (e.g., recovery after a disaster, political views).

Meta-Analytic Procedures

Following the conventions of meta-analysis, for each study we identified the most comprehensive model predicting each outcome (e.g., if a study examined both self-reported aggression and arrest for a violent crime, we identified the most comprehensive model predicting each, amounting to two tests) and treated each measure of disorder in the model as a separate test. A single study then generated one or more records, with each record describing a unique test of the relationship between an outcome and a measure of disorder. The final database included $n = 229$ studies with $k = 721$ tests.

For each test, we recorded estimated effect size and standard error (or estimated standard error based on exact p -value); a list of all covariates in the model; and whether each test estimated a main or interaction effect. Measures of disorder were coded as reflecting physical disorder, social disorder, or a combination, and as utilizing objective audits or resident perceptions. We further subcategorized studies that used resident perceptions by whether they used the perceptions of the focal individual or the aggregation of many resident perceptions. We coded covariates included in each test to determine whether it accounted for one or more measures of affluence (e.g., income) and social process (e.g., collective efficacy). At the study level, we coded aspects of methodology, including characteristics of the sample and nature of the design (longitudinal versus cross-sectional versus experimental).

We limited the meta-analysis reported here to models testing outcomes categorized as (a) socio-behavioral outcomes, which were further limited to offending behavior, aggression, or both ($n = 19$), or (b) neighborhood perceptions ($n = 77$). We also limited the meta-analysis to tests of main effects (excluding interaction effects) on neighborhood residents (excluding studies of visitors to a neighborhood or experiments exposing naïve individuals to unfamiliar neighborhoods).

In order to compare effect sizes of different types (e.g., standardized and unstandardized betas, odds ratios), we used standard errors to calculate t -values and then converted them to Pearson's r values (using the `compute.es` package in R) (Del Re 2015). Many studies tested multiple outcomes within a given category, but we could not treat them as independent cases because they utilized the same database. For this reason, we calculated the average effect size for all models in a study testing an outcome within the category of interest, facilitating a study-level analysis. Because tests within a study generally used the same set of disorder measures and covariates, we were able to generalize these to the study level as the maximum value for all tests in that study (i.e., if any test used a particular disorder measure or covariate, the study as a whole was credited for it).

We used linear weighted mixed-effects models that treated each study-level average effect size as a single observation and weighted its contribution to the model using its variance (using the `metafor` package in R) (Viechtbauer 2015). This approach permitted tests of both global effect size and moderator variables (i.e., different methodological features of studies). Before running models, we omitted 2 studies examining aggression, 12 studies examining fear of crime, and 8 studies examining neighborhood attachment because they did not contain sufficient information to be included in the meta-analysis (i.e., missing standard error) and this information could not be collected from the study authors. Following outlier analyses, we omitted one additional study examining aggression that had undue influence on the model [Cook's $d > 4 * \text{mean}(d)$]; additionally, the study utilized a scale of social disorder that included items about general level of crime, potentially confounding its relationship with the outcome measure, for which we exclude it from all subsequent analysis. We identified no studies in the neighborhood perception categories as outliers.

RESULTS

This section reports the results, beginning with the details of the included studies, followed by the meta-analysis testing each of the pathways proposed by BWT.

Details of Included Studies

Searches of the literature identified 18 studies that tested disorder's effect on at least one outcome reflecting aggressive behavior or offending (totaling 39 unique models; e.g., youth violence, intimate partner violence) and 77 studies that tested disorder's effect on at least one outcome reflecting a perception of or attitude toward neighborhood conditions (totaling 267 unique models; e.g., fear

Table 1 Number of studies and their measurement methodologies for tests of the effects of disorder on aggression and neighborhood perceptions

	Aggression	Neighborhood perceptions
Total number of studies	18	77
Measures of perceived disorder (aggregate measure)	10 (2)	69 (13)
Measures of objective disorder	8	18
Measures of physical disorder	6	34
Measures of social disorder	14	65
Control for social processes	6	36
Control for socioeconomic status	13	57

of crime, neighborhood satisfaction, attachment). We classified this latter set of studies into four subgroups based on the outcomes tested: fear of crime ($k = 56$), neighborhood attachment or satisfaction ($k = 18$; hereafter, neighborhood attachment), perceptions of neighborhood social dynamics ($k = 10$), and perceptions of disorder (treated as an outcome; $k = 6$).³ We focus on the first two subgroups as they are most relevant to our question of interest and are not confounded with our other methodological questions. A full list of studies testing offending or aggressive behaviors, fear of crime, or neighborhood attachment is provided in the **Supplemental Materials, Appendix A**.

Methodologies for the measurement of disorder were largely segregated across studies, particularly regarding the use of perceived or objective measures (see **Table 1**). No study analyzing aggression measured disorder in both ways, and only ten studies (13%) examining neighborhood perceptions did. A similar segregation was visible for the utilization of measures of social or physical disorder; two studies examining aggression (11%), and 22 studies examining neighborhood perceptions (29%) had both. Notably, 48 studies (64%) that examined neighborhood perceptions and attitudes measured disorder exclusively through a survey scale of the focal individual's own perceptions of the neighborhood's level of disorder. In other words, in testing the effect of disorder on a focal individual these studies failed to utilize any independent measure drawn from either objective observations or the aggregate of multiple resident perceptions. In terms of major correlates, 6 of the 19 studies (32%) testing the effect of disorder on offending or aggression controlled for a form of social process (e.g., collective efficacy, social cohesion), and 36 studies (47%) on neighborhood perceptions did the same. All but 5 (68%) of the studies examining offending or aggression controlled for socioeconomic status, as did 57 studies (71%) assessing neighborhood perceptions.

Meta-Analysis

The results of the meta-analysis are organized by disorder's effect on (a) aggression and (b) neighborhood perceptions.

Disorder's effect on aggression. The global effect of disorder on aggression was positive and significant but quite small ($r = 0.007$, $se = 0.002$; $p < 0.001$), and the model exhibited sufficient heterogeneity to justify tests of moderators ($Q_{df=14} = 58.83$, $p < 0.001$). An initial test found

³The four groups sum to more than 77 because some studies tested more than one type of outcome.

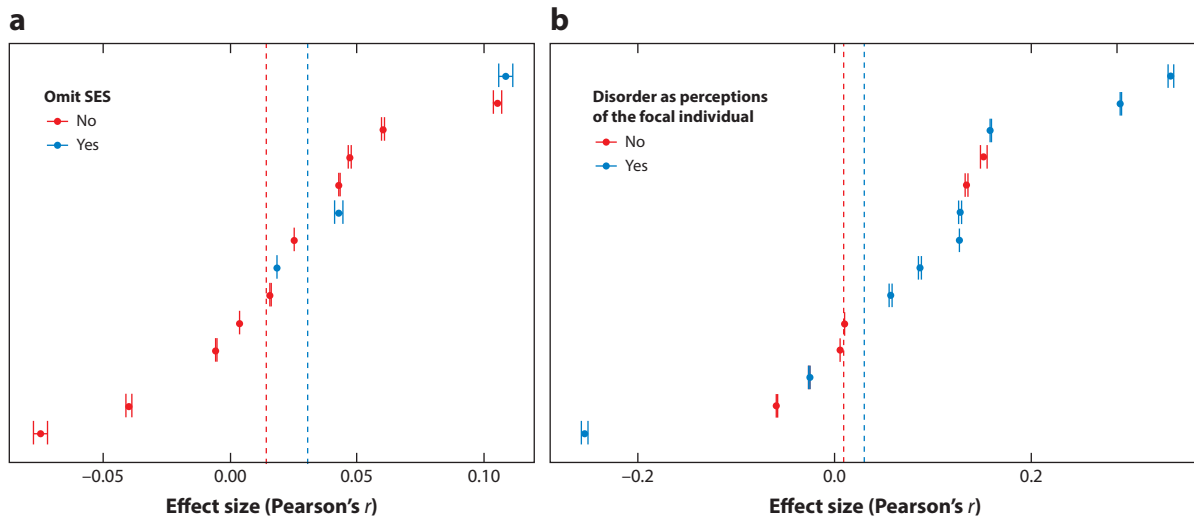


Figure 1

Forest plots of the distribution of estimated effects (with standard errors) across studies of disorder and the global estimated effect (*dashed lines*). (a) Effect on aggression, comparing studies that do and do not omit measures of socioeconomic status (SES). (b) Effect on neighborhood attachment, comparing studies that do and do not measure disorder as the perceptions of the focal individual.

no meaningful difference in the predicted effect of disorder on aggression based on whether it was measured as perceived versus objective [perceived: $r = 0.01$, $se = 0.02$, *ns* (not significant)], physical versus social disorder (social: $r = -.01$, $se = 0.02$, *ns*), or their interaction ($r = 0.005$, $se = 0.02$, *ns*). Because of the small number of studies, we assessed the effect of the omission of important control variables independently. Omitting measures of socioeconomic status sizably increased the estimated effect size ($r = 0.02$, $se = 0.01$, $p < 0.05$), as illustrated in **Figure 1a**. Omitting social process had no significant effect ($r = -.01$, $se = 0.01$). This model also found that, when controlling for these variables, the effect of disorder on aggression was no longer significant ($r = 0.01$, $se = 0.01$, *ns*). See **Table 2** for all model details.

A series of sensitivity analyses evaluated the robustness of these findings by using only studies that were conducted outside the United States; utilized multilevel models or other more robust methodologies for disentangling the effects of neighborhood conditions from individual-level characteristics; and analyzed juveniles instead of or in addition to adults, as the two groups might respond differently to contextual factors (see **Supplemental Materials, Appendix B** for full results). Results from the sensitivity analyses were largely consistent with the global analysis, although in a few cases the estimates became more precise, meaning that parameters were significant without actually growing in magnitude. This was true for aggression in juveniles even when accounting for social process and socioeconomic status ($r = 0.02$, $se = 0.01$, $p < 0.05$). When limiting to studies from the United States or using multilevel models, the inferred effect of disorder on aggression was significant but heavily diminished to a level that is likely practically meaningless (United States only: $r = 0.004$, $se = 0.002$, $p < 0.05$; multilevel models only: $r = 0.004$, $se = 0.001$, $p < 0.05$). In addition, for studies conducted in the United States, perceived disorder did have a significant additional effect on aggression ($r = 0.02$, $se = 0.01$, $p < 0.01$).

Lastly, the majority of studies were from published articles, creating the potential for publication bias. We used the trim-and-fill procedure developed by Duval & Tweedie (2000) to estimate

Supplemental Material >

Table 2 Estimated effects in terms of Pearson’s *r* (with standard errors) of disorder on aggression and neighborhood perceptions and the influence of measurement methodologies on inferred effects, as derived from mixed-effects models^a

Variable	Aggression		Fear of crime		Neighborhood attachment	
			Disorder as perception of individual	Other measures of disorder	Disorder as perception of individual	Other measures of disorder
Intercept	0.02 (0.02)	0.01 (0.01)	−0.02 (0.01)	−0.05 (0.02)	0.03 ^c (0.009)	0.01 (0.01)
Social disorder	−0.01 (0.02)	ND	0.14 ^c (0.01)	0.01 (0.01)	ND	ND
Perceived disorder	0.01 (0.02)	ND	ND	ND	ND	ND
Social perceived ^b	0.005 (0.02)	ND	ND	ND	ND	ND
Omitted social process	ND	−0.01 (0.01)	−0.02 (0.01)	0.08 ^c (0.01)	0.09 ^c (0.01)	−0.01 (0.01)
Omitted socio-economic status	ND	0.02 ^c (0.01)	0.08 ^c (0.01)	0.10 ^c (0.01)	0.07 ^c (0.02)	0.12 ^c (0.03)
<i>N</i> (studies)	15		36	14	9	6

^aAll predictors are dichotomous variables with 1 equal to the variable’s name.

^b*p* < 0.05.

^c*p* < 0.001.

Abbreviation: ND, not determined.

the effect of data censoring on the findings of meta-analyses. This procedure examines the distribution of effect sizes around the mean effect size, whereby the y-axis represents sample size (standard error) and the x-axis represents effect size. In the absence of publication bias, effect sizes should be evenly distributed around the mean effect (Wilson 2009). The analysis indicated that four studies were missing from the left side of the distribution (i.e., negative effect of disorder on aggression), although this had negligible impact on the estimated true effect ($r = 0.005$, $se = 0.002$, $p < 0.05$).

Disorder’s effect on neighborhood perceptions. Similar to the analysis of aggression, the estimated global effect of disorder on the two main types of neighborhood perceptions was significant ($r = 0.06$, $se = 0.002$, $p < 0.001$; with neighborhood attachment reverse-coded; i.e., a positive effect indicates greater fear of crime and lower attachment). The estimated effects were identical between the two subcategories (fear of crime: $r = 0.06$, $se = 0.002$, $p < 0.001$; neighborhood attachment: $r = 0.06$, $se = 0.003$, $p < 0.001$), although they featured different levels of data censoring. Trim-and-fill models found that an estimated 18 studies were missing from the left side of the distribution for fear of crime (adjusted estimated effect size: $r = 0.02$, $se = 0.002$, $p < 0.001$), but that there was no observed publication bias for neighborhood attachment.

It is possible that the inferred relationship between disorder and neighborhood perceptions are confounded with individual-level measures of perceived disorder, inflating support for BWT. Because many studies utilize this measurement technique in conjunction with other measures of disorder, we reran the analyses twice: once for models that tested the effect of disorder measured as the perceptions of the focal individual and once for all other models (including those that used multiple residents’ perceptions to assess disorder as well as objective audits; see **Table 2** for models). If a study utilized both of these two measurement schemes, the models for each scheme were included in the appropriate sub-analysis.

When disorder was measured through the perceptions of the focal individual, it predicted lower neighborhood attachment ($r = 0.03$, $se = 0.009$, $p < 0.001$) and greater fear of crime, but the latter was true only for social disorder (intercept: $r = -.02$, $se = 0.01$, ns ; social disorder: $r = 0.14$, $se = 0.01$, $p < 0.001$). A similar comparison between measures of disorder was not possible for neighborhood attachment, as all studies except one measured social disorder. Measures of disorder were no longer predictive when they were independent of the focal individual's perceptions (fear of crime, intercept: $r = -.05$, $se = 0.02$, ns ; fear of crime, social versus physical disorder: $r = 0.01$, $se = 0.01$, ns ; neighborhood attachment: $r = 0.01$, $se = 0.01$, ns). In fact, the direction of the effect on fear of crime was the opposite of that posited by BWT. **Figure 1b** captures visually this stark distinction between the inferences derived from the two methodological approaches.

The effect of omitting measures of socioeconomic status and social process consistently inflated estimates of disorder's effects. Omitting socioeconomic status predicted higher estimates for the effect of disorder on both fear of crime and neighborhood attachment, regardless of how disorder was measured (fear of crime, perceptions of focal individual: $r = 0.08$, $se = 0.001$, $p < 0.001$; neighborhood attachment, perceptions of focal individual: $r = 0.07$, $se = 0.02$, $p < 0.001$; fear of crime, independent of focal individual: $r = 0.10$, $se = 0.01$, $p < 0.001$; neighborhood attachment, independent of focal individual: $r = 0.12$, $se = 0.03$, $p < 0.001$). The omission of social process had the same effect on fear of crime when disorder was measured independently of the focal individual ($r = 0.08$, $se = 0.01$, $p < 0.001$) and on neighborhood attachment when disorder was measured as the perceptions of the focal individual ($r = 0.09$, $se = 0.01$, $p < 0.001$).

Again, we assessed the robustness of our models with a series of sensitivity analyses (see **Supplemental Materials, Appendix B** for full results). We first limited sensitivity analyses to studies that were multilevel and then to studies conducted in the United States (the latter for studies of fear of crime only, as no study examining neighborhood attachment was conducted outside the United States). No studies exclusively analyzed the neighborhood perceptions of juveniles. There were two notable differences in the results. First, limiting sensitivity analyses to only US studies and studies using multilevel models generated a higher estimate of the effect of disorder on fear of crime when measured as the perceptions of focal individual (US studies only: $r = 0.29$, $se = 0.01$, $p < 0.001$; multilevel models: $r = 0.45$, $se = 0.01$, $p < 0.001$). Because nearly all multilevel studies were conducted in the United States, it is difficult to determine which of these factors is more important. On the one hand, it might point to an effect of America's heavy incorporation of BWT into both policy and public discourse. The inflated parameters might also arise from the ability of multilevel models to disentangle individual- and group-level characteristics. In non-multilevel models, the test of a focal individual's perception of disorder confounds the true conditions with the individual's description of them. A multilevel model isolates the focal individual's positive or negative deviation from objective reality, and this best addresses the concern that an individual's assessments of disorder and crime in their neighborhood are one and the same (e.g., Gau & Pratt 2008). Unexpectedly, for US studies the effect of omitting social process became negative for the effect of disorder measured as the perceptions of the focal individual ($r = -.18$, $se = 0.01$, $p < 0.001$), suggesting that such measures might be necessary as controls to reveal the full extent of the relationship between self-reported disorder and fear of crime.

DISCUSSION

In sum, the meta-analysis demonstrated that neighborhood disorder does not have the consistent influence on resident behaviors postulated by BWT. For the first pathway, neither physical nor social disorder in a neighborhood predicted an increase in the aggression of residents when controlling for different methodological approaches, although there may be a modest effect on

aggression in juveniles. For the second pathway, disorder did not predict fear of crime and only predicted neighborhood attachment when measured as the perceptions of the focal individual; even in the latter case, the effect was modest. There was some evidence that disorder, particularly when measured as the perceptions of the focal individual, had a more substantial effect on fear of crime in studies in the United States.

The assessment here contradicts the large number of studies that have claimed to find support for BWT (see Weisburd et al. 2015), but the indication is that such evidence may have been the overstated result of research designs that led to weak or imprecise inferences. In turn, a more robust understanding of the relationship between disorder and behavior will force us to reconceptualize its role in urban neighborhoods. Before proceeding to these two themes, however, it is first necessary to address a few limitations. First, although the number of studies available was large for a meta-analysis, this was in part because we combined outcome variables we deemed to be conceptually similar. In doing so, we assumed that these constructs (e.g., fear of gangs and fear of homicide) are in fact similarly responsive to disorder. Second, readers should not interpret our results as an evaluation of social disorganization theory and the role of social processes in driving neighborhood effects on aggression, fear of crime, or attachment. The models here explicitly examined only whether social processes accounted for some portion of the estimated relationship between disorder and these outcomes and did not test any direct effect of social process. Further, our meta-analysis centered on disorder, leaving out many studies that tested only the effect of collective efficacy or another social process.

Looking Through Broken Windows: Weaknesses in Methodology and Inference

The analysis here has revealed that part of the difficulty surrounding the debate on the behavioral effects of disorder is methodological. Simply put, have researchers been using the proper research designs to test the premises of BWT? Or have they been inadvertently finding evidence for BWT in studies that are actually testing for other behavioral phenomena? Here, we identify three such weaknesses. The first is the attempt to extrapolate the results of experiments to real-world contexts. The second is the failure to recognize the potential confound between perceptions of disorder and an individual's overarching judgment of a neighborhood. The third is the omission of key correlates that might account for part or all of the correlation between disorder and individual outcomes. We discuss these in turn and then address a fourth issue that captures the potentially reinforcing nature of science and public discourse, as demonstrated by the somewhat anomalous relationship between disorder and behavior observed in samples from the United States.

Translating the results of controlled experiments to real-world phenomena often suffers from a lack of ecological validity, i.e., whether the context of the study can be generalized. Experimental work on BWT has been important in revealing psychological mechanisms that are attuned to the presence of disorder, mechanisms that are necessary for the theory to hold at all. However, controlled experiments do not guarantee that those mechanisms are pertinent to daily life. This mismatch could arise in two ways. First, experiments on BWT effects may magnify the salience of disorder for attitudes and behavior. Typically, such studies expose individuals to unfamiliar neighborhoods of varying levels of upkeep (Kotabe et al. 2016, O'Brien et al. 2014b, O'Brien & Wilson 2011, Yang & Pao 2015), or they use nonresidential, pass-through areas to construct disorderly treatments (Cialdini et al. 1990; Keizer et al. 2008, 2013). In contrast to the participants in either of these designs, residents are well acquainted with their neighborhoods. They may have habituated to the environment and are thus less sensitive to elements of disorder in the scenery. Further, experiments are designed to direct participants' attention specifically to variations in disorder, whereas residents have access to potentially more useful information about the neighborhood and

its social context, knowledge that will more directly influence their own decisions while there (O'Brien & Kauffman 2013).

A second issue pertaining to ecological validity is the question of cross-norm disinhibition discussed above. Experiments have found that the observation of one type of violation might encourage individuals to commit other low-level violations (Keizer et al. 2008), but disinhibition implies that an individual needs to have an initial inclination toward a given behavior for disorder to encourage it. In this light, it is a greater leap to assume that the observation of disorder leads to more serious violations, like violent crime, which are relatively uncommon acts and generally occur under distinctive social circumstances. As such, it is not sufficient to point to studies demonstrating disorder's ability to elicit low-level violations as evidence that it has the same effect for violent crime (e.g., Kelling & Bratton 2015). It appears instead that they are insufficient to drive the emergence of crime as a response to disorder.

The second methodological concern regards the pathway from disorder to neighborhood attitudes. This relationship appears to depend largely on people's tendency to use disorder to make wholesale judgments of the quality of one's neighborhood (O'Brien et al. 2014b, O'Brien & Wilson 2011). Previous survey work, for example, has found it statistically impossible to disentangle respondents' perceptions of disorder and crime (Gau & Pratt 2008, 2010). This might not be such a problem if objective disorder and perceived disorder were equivalent, but they are not. It is well-known that there is considerable within-neighborhood variation in accounts of local levels of disorder and that these assessments depend only partially on disorder, also drawing on other neighborhood characteristics, such as racial composition (Franzini et al. 2008, Sampson & Raudenbush 2004). Consequently, there is a loose relationship between objective and perceived disorder, leading to our own finding that objective disorder has little to no consistent effect on behavior. It is then concerning to note the extent to which research on BWT has relied on measures that inherently confound disorder with neighborhood judgments. Nearly two-thirds (64%) of studies examining neighborhood perceptions measured disorder exclusively through the perceptions of the same focal individual whose outcomes were being analyzed. Although such measures are less expensive to collect, they create a systematic weakness in the evaluation of BWT that has likely inflated the inferred relationship between disorder and fear of crime and neighborhood attachment.

The third and final methodological concern regards the inclusion of control variables in tests of BWT. Here, we focused on socioeconomic status and social processes, both of which are deeply intertwined with disorder, crime, and neighborhood perceptions. The former is considered a fundamental descriptor of urban neighborhoods that should be included as a control variable in any study of urban neighborhoods, yet 20% of studies examining aggression and approximately half of studies examining neighborhood perceptions did not do so. It turned out that such studies consistently overestimated the effects of disorder on behavior. Similarly, social processes were included in only 44% of all studies, inflating estimates of the effect of disorder on neighborhood perceptions. To be fair, the concept of collective efficacy, which repopularized attention to social processes in criminology, was only proposed in 1997 (Sampson et al. 1997) and subsequently articulated as an alternative to BWT in 1999 (Sampson & Raudenbush 1999). However, the proportion of studies accounting for social process is nearly identical if we limit inclusion to studies of neighborhood perceptions published after 1999 (44%). Altogether, these omissions in model specification appear to be leading researchers to overstate the impact of disorder on behavior.

A final concern of note is the divergent results between studies on samples from the United States and those from other countries. Whereas the global analysis consistently found no effect of objective disorder on behavior, limiting inclusion to samples from the United States revealed a modest but significant relationship between perceived disorder and aggression and a heavily inflated effect of disorder as reported by the focal individual; the null effect on neighborhood

attachment was specific to the United States, as no studies examining it were conducted in other countries. These distinctions may be a manifestation of the particular currency that the crime-disorder link holds in both American public policy and popular culture, in which case it would suggest that the broad discussion of BWT is a self-fulfilling prophecy. As BWT has become a mainstay of how Americans discuss crime in urban communities, it is likely to have the very effects that Kelling & Wilson (1982) proposed it to have, especially on fear of crime. It is important to note that most of the other studies here were conducted in other western, industrialized countries, such as Canada, the United Kingdom, and the Netherlands, and thus this is not a result of combining incomparable contexts. If one were to remove the US studies from the sample, the impacts of disorder on behavior in these other locales are even smaller.

Reconceptualizing the Role of Urban Disorder: Theory and Practice

The findings here undermine the basic premise of BWT that disorder alters behavior in a neighborhood in a manner that leads to increases in crime. In turn, it adds to a growing body of literature that calls into question the efficacy of broken windows or zero-tolerance policing practices (Skogan & Frydl 2004). As Weisburd et al. (2015, p. 603) concluded from an in-depth review of broken windows policing studies, any crime reduction observed during such interventions “may be the result of other mechanisms of crime control, such as deterrence, opportunity reduction, or incapacitation” but not through the proposed mechanism of discouraging crime by eliminating low-level misdemeanors. Such insights do not necessarily render disorder irrelevant, but they do force us to reconsider what role it plays in the urban landscape. Here, we examine four alternative interpretations of this role: (a) the social escalation model, which emphasizes private rather than public disorder in instigating crime; (b) disorder’s ability to create ecological advantages that facilitate, but do not necessarily encourage, crime; (c) the pathways by which disorder might still impact public health; and (d) disorder management, or custodianship, as a collective task that reveals how communities function. Importantly, each of these perspectives maintains the importance of disorder for urban neighborhoods: They all articulate causal mechanisms distinct from BWT. Consequently, they call for revised policies and practices that target disorder in a manner that is consistent with its actual impact on the neighborhood.

Social escalation. O’Brien & Sampson (2015) used administrative records to test BWT anew at the neighborhood level. By leveraging calls for government services through the 311 and 911 hotlines in Boston, MA, they were able to track six different dimensions of disorder and violent crime plus homicide rates across multiple years. An important distinction of this approach was that it permitted an examination of not only public disorder but also private disorder in the form of domestic disputes, landlord-tenant troubles, and similar indicators of a conflict between people who share space. This private disorder has never been a part of BWT reasoning and reflects local social dynamics more than signals of criminal threat and opportunity. Longitudinal models found that private disorder was a strong predictor of future increases in violence and gun use within a neighborhood, but that public disorder, whether physical or social, had very little relationship with the same outcomes. Based on these findings, the authors proposed a social escalation model of the relationship between disorder and crime. They argued that certain forms of disorder are evidence of low-level interpersonal conflict that might eventually spill into public spaces or escalate if left unchecked.

The social escalation model is similar to BWT in that it describes a progression from disorder to crime, but it casts disorder’s role in a different way. Traditionally, BWT treats disorder as a signal that elicits and paves the way for crime. In contrast, the social escalation model sees disorder as an

early-stage symptom that, if exacerbated, could erupt in violence. Whereas the former posits that disorder modifies the behaviors of observers, the latter sees the participants in disorder as being themselves at risk for engaging in more serious conflicts. Both call for law enforcement strategies that limit disorder, but the target and approach differ strikingly. BWT tactics are often criticized for harming police-community relations by being too harsh on low-level misdemeanors (Meares 2015). Instead, the amelioration of private social conflict requires a softer, community-oriented touch. It will likely draw more on the use of social workers and others who are able to mitigate such conflicts before they escalate. This is in keeping with recent evidence that policing strategies that prioritize partnerships with community residents and businesses to solve local problems are far more effective for crime prevention (Braga et al. 2015).

Ecological advantages. Although disorder might not elicit crime, there are certain types of disorder that can create ecological advantages for criminal activity. This has been described most evocatively by St. Jean (2007) in his ethnography of a high-crime neighborhood in Chicago. In his interviews with local drug dealers, they rejected the premise of BWT outright, claiming that abandoned lots and houses and other forms of dilapidation simply reflect that the people who live there are poor and do not have the time or resources to maintain the building. Instead, they argue that disorder is helpful for hiding or selling drugs undetected. The dealers described how certain busy, disorder-laden intersections provide sufficient cover for sales, and how abandoned buildings and run-down parts of town are rarely visited by police. Consistent with this story, Furr-Holden et al. (2011) found that marijuana use among adolescents increased in neighborhoods with more abandoned houses, likely because there are unmonitored spaces where drug use can occur undetected.

As with the social escalation model, attention to ecological advantages maintains the importance of disorder in an urban neighborhood but redefines why this is the case. Under this perspective, disorder does not encourage crime but makes it easier to commit crimes. This only matters if people already have some intent to commit said crime in that space, something St. Jean (2007) stresses as he points out that not all places with high disorder in his study had high crime. It also then calls for a different policy approach that specifically targets the types of disorder that enable crime. For example, painting over graffiti will likely be less effective than eliminating the opportunities created by abandoned lots and houses. An initial study by Branas et al. (2016) has borne this out, finding that the proactive remediation of abandoned buildings and vacant lots led to drops in violent crime in the neighborhood. Notably, their study found no effect on nonviolent crime, suggesting the effect was best attributed to ecological advantages for hiding guns, not to any signals that the remediation sent about local maintenance and monitoring.

Disorder and public health. Moving beyond crime, extensions of BWT have argued that neighborhood disorder can impact the public health of residents. Similar to the current study, we might articulate this relationship in terms of a series of hypothesized pathways. There are four such pathways that have been referenced in the literature. First and most simply, disorder can pose a direct threat to health posed by the unsanitary conditions or unstable structures that constitute disorder. This pathway has received policy attention of late (de Leon & Schilling 2017), particularly in terms of asthma (Corburn et al. 2006, Vo et al. 2017) and physical injury (Garvin et al. 2012, Nicklett et al. 2017). However, it is important to note that it is not clear whether such effects result from neighborhood-level context or the residence of a given individual (Holt et al. 2012). In contrast, each of the other three relies on one of the psychological mechanisms for processing information in the neighborhood environment already described here.

The first pathway from disorder to public health outcomes again leverages cross-norm disinhibition, suggesting that the presence of disorder encourages impulsive behaviors in general. The

effect, then, would extend from criminal behavior to risky sexual behavior and alcohol and drug use (Burgard & Lee-Rife 2009, Cohen et al. 2000, Latkin et al. 2007). The second and third pathways from disorder to public health both draw on the second psychological mechanism posited by BWT: that disorder signals danger and heightens fear in local residents. This may impact health in two ways. It can discourage seemingly mundane behaviors that have far-reaching implications for health, like taking walks in one's neighborhood (Heinrich et al. 2007, Kneeshaw-Price et al. 2015, Molnar et al. 2004, Roman et al. 2013). Possibly more insidious, the psycho-social model of disadvantage argues that many of the health disparities experienced by the poor result from the more stressful contexts they inhabit (Massey 2004). Because disorder might signal a threatening environment, it can act as an omnipresent stressor in a neighborhood, thereby increasing the risk for related mental and physical maladies (Browning et al. 2013, Haney 2007, Wen et al. 2006). Some work has even extended these hypotheses to the examination of disorder's relationship with biomarkers associated with allostatic load (i.e., the experience of stress) (Dulin-Keita et al. 2012, Theall et al. 2013).

The relationship observed here between perceived disorder and neighborhood attitudes would support the notion that the experience of a neighborhood can have important effects on mental health and physical activity, even if objective disorder is less important than anticipated. It may also be that the conscious perception of disorder plays a different role for health outcomes, making objective disorder more salient. For example, one need not fear disorder to find it unattractive, and in turn to decide not to go walking in the neighborhood as often. The suggestion that disorder might lead to more risky health behaviors via cross-norm disinhibition, however, would seem to have less empirical support. That said, these are open empirical questions that need to be explored.

Custodianship. As we noted at the outset, more than a century of urban thought had already considered disorder an important element of the urban landscape before Kelling & Wilson (1982) proposed BWT. Early theorists (Booth 1903, Jacobs 1961, Park et al. 1925/1984) did not attribute a causal role to disorder but maintained that it was an important reflection of the effort residents were willing and able to exert for the good of the neighborhood. In the spirit of these seminal perspectives, we might treat custodianship, or the actions that prevent or eliminate disorder (O'Brien 2015), as a model case for understanding how communities accomplish shared goals. Put another way, it allows us to observe how collective efficacy is realized through the actions of local residents. This is not an entirely new line of inquiry, as scholars have traditionally linked efforts by individuals to maintain urban spaces with territoriality (Brown 1987, Brown & Werner 1985, Harris & Brown 1996, Perkins et al. 1993, Taylor 1988, Taylor et al. 1984). However, this previous work reached something of a dead end in the early 1990s, hampered by methodological difficulties. Simply put, custodianship is a sufficiently rare act in that it is time- and cost-prohibitive to conduct an observational protocol that would permit detailed cross-neighborhood comparisons, not to mention analyses of individual differences within those neighborhoods.

The study of custodianship has been rejuvenated by the advent of a new innovation in urban governance: 311 systems. These systems provide one or more channels by which members of the public can request nonemergency government services, typically including a telephone hotline (e.g., 3-1-1) and associated online applications. A sizable proportion of these requests regard issues with public spaces or infrastructure, what we might call the urban commons, such as graffiti, street light outages, and potholes. The systems compile these requests in a work-order queue. This same queue is essentially a database tracking acts of custodianship across a city. In major cities, this database amounts to hundreds of thousands of records a year. Additionally, many cities allow users to create personal accounts by which they can follow their requests, permitting studies that examine individual-level patterns of custodianship both within and between neighborhoods.

O'Brien and colleagues have leveraged this opportunity, conducting a series of studies on the management of disorder and deterioration in urban neighborhoods. This work has found that custodianship is a hyperlocal act, occurring predominantly around an individual's home (O'Brien 2015, 2016a), and has demonstrated that it is indeed motivated by territoriality (O'Brien et al. 2014a, 2017). It also found that individuals often specialize in either natural deterioration such as street light outages or manmade incivilities such as graffiti (O'Brien 2016c). In considering the collective consequences of these outcomes, a later study uncovered a division of labor in which occasional reporters (or typical custodians) and more frequent reporters (or exemplars) complement each other and are thus both necessary to the realization of collective efficacy in the maintenance of the urban commons (O'Brien 2016b).

One of the main themes of research on disorder is how it suggests ways for policymakers and practitioners to address it. Substantively, the research on custodianship via 311 systems suggests ways that city governments might better target their messaging of the system to encourage participation or support neighborhoods given a known profile of custodianship. These findings are particularly practical because they are drawn from the administrative data generated by the system itself, making them immediately applicable to the management of city services. Further, 311 is a relatively recent innovation in urban policy and one that is still evolving and growing in popularity. There is both a desire and need on the part of policymakers to learn how to help it be most effective, emphasizing the value of such research in the way disorder is addressed in cities in the coming years.

CONCLUSION

Our meta-analysis of the effects of neighborhood disorder on the behavior of residents adds to the growing body of work calling into question the theoretical premise of BWT. Rather than just critique a single, albeit prominently influential idea, we have sought to reconsider the relationships among disorder, resident behaviors, and crime. This has far-reaching implications for theory and research and the lines of inquiry that will have purchase moving forward. It also points to the need for critical adjustments in law enforcement strategies that direct attention to disorder, especially in those US cities where BWT has been embraced so strongly. This presents a special opportunity to adopt strategies that are more evidenced by data to be effective while also moving away from tactics that have tended to undermine fairness and legitimacy. In the words of Lewin, "there is nothing as practical as a good theory" (Lewin 1943, p. 118), and as our theories of crime in communities are refined by the scientific process, so must our approaches to law enforcement.

DISCLOSURE STATEMENT

The authors are not aware of any affiliations, memberships, funding, or financial holdings that might be perceived as affecting the objectivity of this review.

LITERATURE CITED

- Booth C. 1903. *Life and Labour of the People in London*. London: Macmillan. 3rd ed.
- Braga AA, Bond BJ. 2008. Policing crime and disorder hot spots: a randomized controlled trial. *Criminology* 46:577–607
- Braga AA, Welsh BC, Schnell C. 2015. Can policing disorder reduce crime? A systematic review and meta-analysis. *J. Res. Crime Delinquency* 52:567–88
- Branas CC, Kondo MC, Murphy SM, South EC, Polsky D, MacDonald JM. 2016. Urban blight remediation as a cost-beneficial solution to firearm violence. *Am. J. Public Health* 106:2158–64

- Bratton W, Knobler P. 2001. *Turnaround*. New York: Random House
- Brown BB. 1987. Territoriality. In *Handbook of Environmental Psychology*, ed. D Stokols, I Altman, pp. 505–31. New York: Wiley
- Brown BB, Werner CM. 1985. Social cohesiveness, territoriality and holiday decorations: the influence of cul-de-sacs. *Environ. Behav.* 17:539–65
- Browning CR, Soller B, Gardner M, Brooks-Gunn J. 2013. “Feeling disorder” as a comparative and contingent process: gender, neighborhood conditions, and adolescent mental health. *J. Health Soc. Behav.* 54:296–314
- Burgard SA, Lee-Rife SM. 2009. Community characteristics, sexual initiation, and condom use among young Black South Africans. *J. Health Soc. Behav.* 50:293–309
- Cialdini RB, Reno RR, Kallgren CA. 1990. A focus theory of normative conduct: recycling the concept of norms to reduce littering in public places. *J. Personal. Soc. Psychol.* 58:1015–26
- Cohen D, Spear S, Scribner R, Kissinger P, Mason K, Widgen J. 2000. “Broken windows” and the risk of gonorrhea. *Am. J. Public Health* 90:230–36
- Corburn J, Osleeb J, Porter M. 2006. Urban asthma and the neighbourhood environment in New York City. *Health Place* 12:167–79
- de Leon E, Schilling J. 2017. *Urban blight and public health*. Rep., Urban Inst., Washington, DC
- Del Re AC. 2015. Package ‘compute.es.’ <https://cran.r-project.org/web/packages/compute.es/compute.es.pdf>
- Dulin-Keita A, Casazza K, Fernandez JR, Goran MI, Gower B. 2012. Do neighbourhoods matter? Neighbourhood disorder and long-term trends in serum cortisol levels. *J. Epidemiol. Community Health* 66:24–29
- Duval S, Tweedie R. 2000. Trim and fill: a simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis. *Biometrics* 56:455–63
- Franzini L, Caughy MOB, Nettles SM, O’Campo P. 2008. Perceptions of disorder: contributions of neighborhood characteristics to subjective perceptions of disorder. *J. Environ. Psychol.* 28:83–93
- Furr-Holden CDM, Lee MH, Milam AJ, Johnson RM, Lee K-S, Ialongo NS. 2011. The growth of neighborhood disorder and marijuana use among urban adolescents: a case for policy and environmental interventions. *J. Stud. Alcohol Drugs* 72:371–79
- Garofalo J, Laub J. 1978. The fear of crime: broadening our perspective. *Victimology* 3:242–53
- Garvin E, Branas C, Keddem S, Sellman J, Cannuscio C. 2012. More than just an eyesore: local insights and solutions on vacant land and urban health. *J. Urban Health* 90:412–26
- Gau JM, Pratt TC. 2008. Broken windows or window dressing? Citizens’ (in)ability to tell the difference between disorder and crime. *Criminol. Public Policy* 7:163–94
- Gau JM, Pratt TC. 2010. Revisiting broken windows theory: examining the sources of the discriminant validity of perceived disorder and crime. *J. Crim. Justice* 38:758–66
- Haney TJ. 2007. “Broken windows” and self-esteem: subjective understandings of neighborhood poverty and disorder. *Soc. Sci. Res.* 36:963–94
- Harcourt BE. 2001. *Illusion of Order: The False Promise of Broken Windows Policing*. Cambridge, MA: Harvard Univ. Press
- Harcourt BE, Ludwig J. 2006. Broken windows: new evidence from New York City and a five-city social experiment. *Univ. Chic. Law Rev.* 73:271–320
- Harris PB, Brown BB. 1996. The home and identity display: interpreting resident territoriality from home exteriors. *J. Environ. Psychol.* 16:187–203
- Heinrich KM, Lee RE, Suminski RR, Regan GR, Reese-Smith JY, et al. 2007. Associations between the built environment and physical activity in public housing residents. *Int. J. Behav. Nutr. Phys. Act.* 4:56
- Holt EW, Theall KP, Rabito FA. 2012. Individual, housing, and neighborhood correlates of asthma among young urban children. *J. Urban Health* 90:116–29
- Hunter A. 1974. *Symbolic Communities*. Chicago: Univ. Chicago Press
- Jacobs J. 1961. *The Death and Life of Great American Cities*. New York: Random House
- Keizer K, Lindenberg S, Steg L. 2008. The spreading of disorder. *Science* 332:1681–85
- Keizer K, Lindenberg S, Steg L. 2013. The importance of demonstratively restoring order. *Public Libr. Sci.* 8:e65137
- Kelling GL, Bratton WJ. 2015. Why we need broken windows policing. *City Journal*, Winter 2015. <https://www.city-journal.org/html/why-we-need-broken-windows-policing-13696.html>

- Kelling GL, Coles CM. 1997. *Fixing Broken Windows: Restoring Order and Reducing Crime in our Communities*. New York: Simon & Schuster
- Kelling GL, Wilson JQ. 1982. Broken windows: the police and neighborhood safety. *Atl. Monthly* 127:29–38
- Kneeshaw-Price SH, Saelens BE, Sallis JF, Frank LD, Grembowski DE, et al. 2015. Neighborhood crime-related safety and its relation to children's physical activity. *J. Urban Health* 92:472–89
- Kotabe HP, Kardan O, Berman MG. 2016. The order of disorder: deconstructing visual disorder and its effect on rule-breaking. *J. Exp. Psychol.* 145:1713–27
- Kubrin CE. 2008. Making order of disorder: a call for conceptual clarity. *Criminol. Public Policy* 7:203–14
- Latkin CA, Curry AD, Hua W, Davey MA. 2007. Direct and indirect associations of neighborhood disorder with drug use and high-risk sexual partners. *Am. J. Prev. Med.* 32:S234–41
- Lewin K. 1943. Psychology and the process of group living. *J. Soc. Psychol.* 17:113–31
- Lipsey MW, Wilson D. 2001. *Practical Meta-Analysis*. Woburn, MA: SAGE
- Markowitz F, Bellair P, Liska A, Jianhong LJ. 2001. Extending social disorganization theory: modeling the relationships between cohesion, disorder, and fear. *Criminology* 39:293–319
- Massey DS. 2004. Segregation and stratification: a biosocial perspective. *Du Bois Rev.* 1:7–25
- Mayhew H. 1862. *London Labor and the London Poor*. London: Griffin Bohn
- Meares T. 2015. Broken windows, neighborhoods, and the legitimacy of law enforcement, or why I fell in and out of love with Zimbardo. *J. Res. Crime Delinquency* 52:609–25
- Milam AJ, Furr-Holden CDM, Leaf PM. 2010. Perceived school and neighborhood safety, neighborhood violence and academic achievement in urban school children. *Urban Rev.* 42:458–67
- Molnar BE, Gortmaker SL, Bull FC, Buka SL. 2004. Unsafe to play? Neighborhood disorder and lack of safety predict reduced physical activity among urban children and adolescents. *Am. J. Health Promot.* 18:378–86
- Nicklett EJ, Lohman MC, Smith ML. 2017. Neighborhood environment and falls among community-dwelling older adults. *Int. J. Environ. Res. Public Health* 14:E175
- O'Brien DT. 2015. Custodians and custodianship in urban neighborhoods: a methodology using reports of public issues received by a city's 311 hotline. *Environ. Behav.* 47:304–27
- O'Brien DT. 2016a. 311 hotlines, territoriality, and the collaborative maintenance of the urban commons: examining the intersection of a coproduction policy and evolved human behavior. *Evol. Behav. Sci.* 10:123–41
- O'Brien DT. 2016b. Lamp lighters and sidewalk smoothers: how individual residents contribute to the maintenance of the urban commons. *Am. J. Community Psychol.* 58:391–409
- O'Brien DT. 2016c. Using small data to interpret big data: 311 reports as individual contributions to informal social control in urban neighborhoods. *Soc. Sci. Res.* 59:83–96
- O'Brien DT, Gordon E, Baldwin-Philippi J. 2014a. Caring about the community, counteracting disorder: 311 reports of public issues as expressions of territoriality. *J. Environ. Psychol.* 40:320–30
- O'Brien DT, Kauffman RA. 2013. Broken windows and low adolescent prosociality: not cause and consequence but co-symptoms of low collective efficacy. *Am. J. Community Psychol.* 51:359–69
- O'Brien DT, Norton C, Cohen J, Wilson DS. 2014b. Local adaptation in community perception: how background impacts judgments of neighborhood safety. *Environ. Behav.* 46:213–40
- O'Brien DT, Offenhuber D, Baldwin-Philippi J, Sands M, Gordon E. 2017. Uncharted territoriality in coproduction: the motivations for 311 reporting. *J. Public Adm. Res. Theory* 27:320–35
- O'Brien DT, Sampson RJ. 2015. Public and private spheres of neighborhood disorder: assessing pathways to violence using large-scale digital records. *J. Res. Crime Delinquency* 52:486–510
- O'Brien DT, Wilson DS. 2011. Community perception: the ability to assess the safety of unfamiliar neighborhoods and respond adaptively. *J. Personal. Soc. Psychol.* 100:606–20
- Park RE, Burgess EW, McKenzie RD. 1925/1984. *The City: Suggestions for Investigation of Human Behavior in the Urban Environment*. Chicago: Univ. Chicago Press
- Perkins DD, Wandersman A, Rich RC, Taylor RB. 1993. The physical environment of street crime: defensible space, territoriality and incivilities. *J. Environ. Psychol.* 13:29–49
- Roman CG, Stodolska M, Yahner J, Shinew K. 2013. Pathways to outdoor recreation, physical activity, and delinquency among urban Latino adolescents. *Ann. Behav. Med.* 45:151–61
- Sampson RJ, Raudenbush SW. 1999. Systematic social observation of public spaces: a new look at disorder in urban neighborhoods. *Am. J. Sociol.* 105:603–51

- Sampson RJ, Raudenbush SW. 2004. Seeing disorder: neighborhood stigma and the social construction of “broken windows.” *Soc. Psychol. Q.* 67:317–42
- Sampson RJ, Raudenbush SW, Earls F. 1997. Neighborhoods and violent crime: a multilevel study of collective efficacy. *Science* 277:918–24
- Skogan W. 1986. Fear of crime and neighborhood change. In *Communities and Crime*, ed. J Albert, J Reiss, M Tonry, pp. 203–29. Chicago: Univ. Chicago Press
- Skogan W. 1990. *Disorder and Decline: Crime and the Spiral of Decay in American Cities*. Berkeley: Univ. Calif. Press
- Skogan WG. 1992. *Disorder and Decline*. Berkeley, CA: University of California Press
- Skogan WG, Frydl K, eds. 2004. *Fairness and Effectiveness in Policing: The Evidence*. Washington, DC: Natl. Acad. Press
- Steenbeek W, Hipp JR. 2011. A longitudinal test of social disorganization theory: feedback effects among cohesion, social control, and disorder. *Criminology* 49:833–71
- St. Jean PKB. 2007. *Pockets of Crime: Broken Windows, Collective Efficacy, and the Criminal Point of View*. Chicago: Univ. Chicago Press
- Taylor B, Koper CS, Woods DJ. 2011. A randomized controlled trial of different policing strategies at hot spots of violent crime. *J. Exp. Criminol.* 7:149–81
- Taylor RB. 1988. *Human Territorial Functioning: An Empirical, Evolutionary Perspective on Individual and Small Group Territorial Cognitions, Behaviors and Consequences*. Cambridge: Cambridge Univ. Press
- Taylor RB. 2001. *Breaking Away from Broken Windows: Baltimore Neighborhoods and the Nationwide Fight Against Crime, Grime, Fear, and Decline*. Boulder, CO: Westview Press
- Taylor RB. 2005. The incivilities or “broken windows” thesis. In *Encyclopedia of Law Enforcement*, ed. LE Sullivan, MS Rosen, DM Schulz, MR Haberland, pp. 31–35. Thousand Oaks, CA: SAGE
- Taylor RB, Gottfredson SD, Bower S. 1984. Block crime and fear: defensible space, local social ties, and territorial functioning. *J. Res. Crime Delinquency* 21:303–31
- Theall KP, Brett ZH, Shirtcliff EA, Dunn EC, Drury SS. 2013. Neighborhood disorder and telomeres: connecting children’s exposure to community level stress and cellular response. *Soc. Sci. Med.* 85:50–58
- Viechtbauer W. 2015. *Metafor: A Meta-Analysis Package for R*. <http://www.metafor-project.org/doku.php>
- Vo P, Bair-Merritt M, Camargo J, Carlos A, Eisenberg S, Long W. 2017. Individual factors, neighborhood social context and asthma at age 5 years. *J. Asthma* 54:265–72
- Weisburd D, Hinkle JC, Braga AA, Wooditch A. 2015. Can policing disorder reduce crime? A systematic review and meta-analysis. *J. Res. Crime Delinquency* 52:567–88
- Weisburd D, Hinkle JC, Famega C, Ready J. 2011. The possible “backfire” effects of hot spots policing: an experimental assessment of impacts on legitimacy, fear and collective efficacy. *J. Exp. Criminol.* 7:297–320
- Wen M, Hawkey LC, Cacioppo JT. 2006. Objective and perceived neighborhood environment, individual SES and psychosocial factors, and self-rated health: an analysis of older adults in Cook County, Illinois. *Soc. Sci. Med.* 63:2575–90
- Wilson DB. 2009. Missing a critical piece of the pie: simple document search strategies inadequate for systematic reviews. *J. Exp. Criminol.* 5:429–40
- Wilson JQ. 1975. *Thinking About Crime*. New York: Basic Books
- Yang S-M, Pao C-C. 2015. Do we “see” the same thing? An experimental look into the black box of disorder perception. *J. Res. Crime Delinquency* 52:534–66
- Zimbardo PG. 1969. The human choice: individuation, reason, and order versus deindividuation, impulse, and chaos. In *Nebraska Symposium on Motivation 1969*, ed. WT Arnold, D Levine, pp. 237–307. Lincoln, NE: Univ. Neb. Press