Authors

Anthony A. Braga, Ph.D.
Professor
Rutgers University
School of Criminal Justice
123 Washington Street
Newark, New Jersey 07102 USA

&
Senior Research Fellow
Harvard University
Kennedy School of Government
79 John F. Kennedy Street
Cambridge, Massachusetts 02138 USA
Anthony_Braga@harvard.edu, braga@andromeda.rutgers.edu

Brandon C. Welsh, Ph.D.
Professor
School of Criminology and Criminal Justice
Northeastern University
Churchill Hall, 360 Huntington Avenue
Boston, Massachusetts 02115 USA

&
Senior Research Fellow
Netherlands Institute for the Study of Crime and Law Enforcement (NSCR)
Amsterdam, The Netherlands
b.welsh@neu.edu

Contact Person

Anthony A. Braga, Ph.D.

Date

First Submission: January 7, 2012
Second Submission: February 10, 2012
Third Submission: June 6, 2012
Sources of Support

Internal Funding

The School of Criminal Justice at Rutgers University will be an intramural source of support for this project. Resources of the Gottfredson Criminal Justice Library will be used to conduct the search for eligible studies and information retrieval.

External funding

We will seek support for the research from external sources such as private foundations and government grant-making agencies.

Background for the Review

Crime policy scholars, primarily James Q. Wilson and George L. Kelling, and practitioners, such as Los Angeles Police Chief William J. Bratton, have argued for years that when police pay attention to minor offenses—such as aggressive panhandling, prostitution, and graffiti—they can reduce fear, strengthen communities, and prevent serious crime (Bratton & Kelling, 2006; Wilson & Kelling, 1982). Spurred by claims of large declines in serious crime after the approach was adopted in New York City in the early 1990s, dealing with physical and social disorder, or “fixing broken windows,” has become a central element of crime prevention strategies adopted by many American police departments (Kelling & Coles, 1996; Sousa & Kelling, 2006).

In their seminal “broken windows” article, Wilson and Kelling (1982) argue that social incivilities (e.g., loitering, public drinking, and prostitution) and physical incivilities (e.g., vacant lots, trash, and abandoned buildings) cause residents and workers in a neighborhood to be fearful. Fear causes many stable families to move out of the neighborhood and the remaining residents isolate themselves and avoid others. Anonymity increases and the level of informal social control decreases. The lack of control and escalating disorder attracts more potential offenders to the area and this increases serious criminal behavior. Wilson and Kelling (1982) argued that serious crime developed because the police and citizens did not work together to prevent urban decay and social disorder.

The available research evidence on the connections between disorder and more serious crime is mixed. In the Netherlands, Keitzer et al. (2008) conducted six field experiments examining the links between disorder and more serious crime and concluded that dealing with disorderly conditions was an important intervention to halt the spread of further crime and disorder. Skogan’s (1990) survey research found disorder to be significantly correlated with perceived crime problems in a neighborhood even after controlling for the population’s poverty, stability, and racial composition. Further, Skogan’s (1990) analysis of robbery victimization data from thirty neighborhoods found that economic and social factors’ links to crime were indirect and mediated through disorder. In his reanalysis of
the Skogan data, Harcourt (1998, 2001) removed several neighborhoods with very strong
disorder-crime connections from Newark, New Jersey, and reported no significant
relationship between disorder and more serious crime in the remaining neighborhoods.
Eck and Maguire (2006) suggest that Harcourt’s analyses do not disprove Skogan’s
results; rather his analyses simply document that the data are sensitive to outliers.
Indeed, the removal of different neighborhoods from Harcourt’s analysis may have
strengthened the disorder-crime connection (Eck & Maguire, 2006).

In his longitudinal analysis of Baltimore neighborhoods, Taylor (2001) finds some
support that disorderly conditions lead to more serious crime. However, these results
varied according to types of disorder and types of crime. Taylor (2001) suggests that
other indicators, such as initial neighborhood status, are more consistent predictors of
later serious crimes. Using systematic social observation data to capture social and
physical incivilities on the streets of Chicago, Sampson and Raudenbush (1999) found
that, with the exception of robbery, public disorder was not significantly related to most
forms of serious crime when neighborhood characteristics such as poverty, stability, race,
and collective efficacy were considered. Sampson and Raudenbush’s findings have been
criticized because their social observation data on disorder were collected during the day
rather than at night (Sousa & Kelling, 2006) and based on their decision to test a model in
which disorder mediates the effects of neighborhoods characteristics on crime rather than
neighborhood characteristics mediating the effects of disorder on crime (Jang & Johnson,
2001). In another analysis, Xu et al. (2005) point out that Sampson and Raudenbush
(1999)’s results actually are supportive of broken windows theory.

The scientific research evidence on the crime control effectiveness of broad-based broken
windows policing strategies, such as quality-of-life programs and order maintenance
enforcement practices, is also mixed. However, there seems to be more research evidence
supporting the crime prevention value of broken windows policing strategies than
refuting it. The New York City Police Department (NYPD) provides the best known
evidence of a macro policy of order maintenance, as it is well documented that officers
were more aggressive in making arrests for minor offenses (Sousa & Kelling, 2006).
Using misdemeanor arrests as a proxy for order maintenance activities, Kelling and
Sousa (2001) found that the NYPD strategy was associated with a significant reduction in
violent crime in the 1990s, after controlling for economic, demographic, and drug use
variables. A similar analysis by Corman and Mocan (2002) found that increased
misdemeanor arrests in New York City during the 1990s had a significant impact on
robbery and motor vehicle theft controlling for economic and criminal justice factors.

Two more recent studies of the impact of order-maintenance policing in New York City
support the idea that policing disorder prevents more serious crime. Rosenfeld,
Fornango, and Renfigo (2007) analyzed the effects of order-maintenance arrests on
precinct-level robbery and homicide trends in New York City between 1988 and 2001,
and concluded that the approach generated small but significant crime reduction gains.
Using a different analytic approach, Messner and his colleagues (2007) analyzed
homicide trends in 74 New York City police precincts between 1990 and 1999, and found
that misdemeanor arrests generated significant reductions in total homicide rates with the
largest impacts on gun homicide rates. This is consistent with Fagan, Zimring, and Kim’s (1998) observation that the kinds of changes in policing associated with broken windows policing might be effective, in part, by taking more guns off the streets through increased police-citizen contacts. More recently, Zimring (2012) suggests that the police action in reducing crime in New York City during the 1990s was not the broken windows policing described by Kelling and Coles (1996) rather it more closely resembled a tight police focus on crime hot spots.

There are also policy evaluations implemented in other jurisdictions that support the perspective that dealing with disorderly conditions generates crime control gains. Two separate randomized controlled trials of disorder policing strategies implemented within a problem-oriented policing framework found the strategy resulted in significant reductions in calls for service to the police in Jersey City, New Jersey (Braga et al., 1999) and Lowell, Massachusetts (Braga & Bond, 2008). The Safer City Initiative, an intervention launched by the Los Angeles Police Department to reduce homeless-related crimes by addressing disorderly conditions associated with homeless encampments, generated modest reductions in violent, property, and nuisance street crimes (Berk & MacDonald, 2010). Other macro-level analyses have generated results supportive of broad-based policing disorder strategies. In California, controlling for demographic, economic, and deterrence variables, a county-level analysis revealed that increases in misdemeanor arrests associated with significant decreases in felony property offenses (Worrall, 2002). Finally, an analysis of robbery rates in 156 American cities revealed that aggressive policing of disorderly conduct and driving under the influence reduces robbery (Sampson & Cohen, 1988).

Many observers, however, argue that it is very difficult to credit a generalized order maintenance strategy with the crime drop in New York in the 1990s. The NYPD implemented the broken windows strategy within a larger set of organizational changes framed by the Compstat management accountability structure for allocating police resources (Silverman, 1999). As such, it is difficult to establish the independent effects of broken windows policing relative to other strategies implemented as part of the Compstat process (Weisburd et al., 2003). Other scholars suggest that a number of rival causal factors, such as the decline in the city’s crack epidemic, played a more important role in the crime drop (Blumstein, 1995; Bowling, 1999). Some academics have argued that the crime rate was already declining in New York before the implementation of any of the post-1993 police reforms, and that New York’s decline in homicide rates were not significantly different from declines experienced in surrounding states and in other large cities that did not implement aggressive enforcement policies during that time period (Karmen, 2000; Eck & Maguire, 2006).

Other evaluations have not found significant crime prevention gains associated with broad-based policing disorder strategies. A recent reanalysis of the Kelling and Sousa (2001) data did not find that a generalized broken windows strategy, as measured by increased misdemeanor arrests, yielded significant reductions in serious crimes in New York City between 1989 and 1998 (Harcourt & Ludwig, 2006). A quasi-experimental evaluation of a quality-of-life policing initiative focused on social and physical disorder
in four target zones in Chandler, Arizona did not find any significant reductions in serious crime associated with the strategy (Katz et al., 2001). An evaluation of a one-month police enforcement effort to reduce alcohol and traffic-related offenses in a community in a Midwestern city did not find any significant reductions in robbery or burglary in the targeted area (Novak et al., 1999). Similarly, a randomized controlled experiment of broken windows policing in three towns in California (Redlands, Colton, and Ontario) found no significant effects on fear of crime, police legitimacy, collective efficacy, or perceptions of crime and social disorder (Weisburd et al., 2011).

Given the mixed policy evaluation findings, a systematic review of the existing empirical evidence is warranted.

Objectives for the Review

This review will synthesize the existing published and non-published empirical evidence on the effects of broken windows policing interventions and will provide a systematic assessment of the crime reduction value of broken windows policing in neighborhoods. It is anticipated that this review will help inform policy makers and police department decision makers regarding the continued use of broken windows policing interventions to reduce crime in neighborhoods. Many police agencies in the United States, United Kingdom, Australia, and other nations currently use broken windows policing as a core crime control strategy and a critical examination of the existing evidence is warranted.

Methods

Criteria for Inclusion and Exclusion of Studies in the Review

Types of Interventions

The general idea of dealing with disorderly conditions to prevent crime is present in myriad police strategies, ranging from “order maintenance” and “zero-tolerance,” where the police attempt to impose order through strict enforcement, to “community” and “problem-oriented policing” strategies where police attempt to produce order and reduce crime through cooperation with community members and by addressing specific recurring problems (Cordner, 1998; Eck & Maguire, 2006; Skogan, 2006; Skogan et al., 1999). While its application can vary within and across police departments, broken windows policing to prevent crime is now a common crime control strategy.

We will consider all policing programs that attempt to reduce crime through addressing physical disorder (vacant lots, abandoned buildings, graffiti, etc.) and social disorder (public drinking, prostitution, loitering, etc.) in neighborhood areas. These interventions will be compared to other police crime reduction efforts that do not attempt to reduce crime through reducing disorderly conditions such as traditional policing (i.e., regular
levels of patrol, ad-hoc investigations, etc.) or problem-oriented policing programs focused on other types of local dynamics and situations.

Types of Areas

Based on the selected literature review above, we expect that our research strategy will yield a diverse set of targeted areas across the identified policing disorder studies. For example, evaluations of broken windows policing strategies in New York City analyzed the citywide effects of the strategy at different units of analysis such as police precincts and police boroughs (Kelling & Sousa, 2001; Corman & Mocan, 2002; Harcourt & Ludwig, 2006; Rosenfeld et al., 2007; Messner et al., 2007). In Los Angeles, evaluators compared crime trends in one policing disorder treatment police division area relative to crime trends in four adjacent police division areas (Berk & MacDonald, 2010). In the Jersey City and Lowell randomized controlled trials, the units of analysis were crime “hot spots” comprised of street block faces and street intersections (Braga et al., 1999; Braga & Bond, 2008).

All area-level studies will be included in our systematic review. Eligible areas can range from small places (such as hot spots comprised of clusters of street segments or addresses) to police defined areas (such as districts, precincts, sectors, or beats) to larger neighborhood units (such as census tracts or a researcher-defined area). It is important to note that this heterogeneity in the units of analysis across studies could have varying and policy-relevant effects on crime prevention outcomes associated with the policing disorder strategies. As such, we will also classify the types of areas to ensure that the review is measuring similar findings across the potentially diverse set of locations subjected to treatment.

Types of Studies

Studies that use comparison group designs, such as randomized controlled trials and quasi-experimental designs (Shadish, Cook, & Campbell, 2002), will be eligible for the main analyses of this review. Only the most rigorous quasi-experimental designs will be included, with the minimum design involving before and after measures of crime in experimental and comparable control areas. In many controlled policing disorder evaluations (e.g. Berk & MacDonald, 2010; Katz et al., 2001), the control group experiences routine modern police responses to crime. Control areas usually experience a blend of traditional police responses (e.g., random patrol, rapid response, and ad-hoc investigations) and opportunistic community problem-solving responses. While disorder interventions developed from community policing initiatives may be present in certain control areas, none of the control areas can engage disorder policing strategies as their main approach to address crime problems.

Types of Outcomes

Eligible studies will have to measure the effects of the broken windows policing intervention on officially recorded levels of crime in areas such as crime incident reports,
citizen emergency calls for service, and arrest data. Other outcomes measures such as survey, interview, systematic observations of social disorder (such as loitering, public drinking, and the solicitation of prostitution), systematic observations of physical disorder (such as trash, broken windows, graffiti, abandoned homes, and vacant lots), and victimization measures used by eligible studies to measure program effectiveness will also be coded and analyzed.

Since area-level studies will be included in this review, particular attention will be paid to studies that measured crime displacement effects and diffusion of crime control benefit effects. Policing strategies focused on specific locations have been criticized as resulting in displacement (see Reppetto, 1976). More recently, academics have observed that crime prevention programs may result in the complete opposite of displacement—that crime control benefits were greater than expected and “spill over” into places beyond the target areas (Clarke & Weisburd, 1994). The quality of the methodologies used to measure displacement and diffusion effects, as well as the types of displacement (spatial, temporal, target, modus operandi) examined, will be assessed.

While all eligible studies must include a crime outcome measure, we will specifically collect data on community satisfaction measures such as citizen attitudes towards police, fear of crime, and other outcomes. Questions have been raised about the legitimacy of specific tactics used by the police to control disorder (Tyler & Fagan, 2008). Inappropriate policing disorder strategies, such as the indiscriminate aggressive tactics used in zero-tolerance approaches, could have negative impacts on police–community relationships. For instance, the heightened use of arrests for minor crimes, such as public drinking and smoking marijuana in plain view, in the NYPD’s order maintenance policing strategies have been criticized as exacerbating already poor relationships between the police and minority communities and increasing citizen complaints about police misconduct and abuse of force (Golub et al., 2007; Greene, 1999; Harcourt & Ludwig, 2007).

**Timeframe**

The 1982 article by Wilson and Kelling will mark the beginning of the timeframe for the inclusion of studies in this review.

**Geographic Scope**

There are no restrictions on the geographical origin of studies for inclusion in this review. The search strategy is international in scope and is not limited to studies in the English language.

**Search Strategies for Identification of Studies**

To identify the studies meeting the criteria of this review, the following nine search strategies will be used:
1. A keyword search will be performed on an array of online abstract databases (see lists of keywords and databases below).

2. The bibliographies of past narrative and empirical reviews of literature that examined the effectiveness of police crime control programs will be reviewed (e.g., Braga, 2008; Eck & Maguire, 2006; Sherman, 1997, 2002; Skogan & Frydl, 2004; Weisburd & Eck, 2004).

3. Forward searches for works that cited seminal broken windows policing studies will be performed (e.g., Wilson & Kelling, 1982; Kelling & Coles, 1996; Sousa & Kelling, 2001).

4. Bibliographies of past completed Campbell systematic reviews of police crime control efforts will be searched (Mazerolle et al., 2007; Weisburd et al., 2008; Bowers et al., 2010; Braga, 2007).

5. Hand searches of leading journals in the field will be performed.¹

6. Searches of policing literature available on governmental and non-profit organization web pages.²

7. Keyword searches of grey literature databases (e.g. Gottfredson Library criminal justice grey literature database maintained by Phyllis Schultze; the System for Grey Literature in Europe, http://www.opengrey.eu/).

8. Keyword searches of abstracts of papers presented at professional criminology and criminal justice conferences (e.g. American Society of Criminology, Academy of Criminal Justice Sciences, International Society of Criminology, U.S. National Institute of Justice Research and Evaluation Conference).

9. After finishing the above searches and reviewing the studies as described later, the list of studies meeting our eligibility criteria will be emailed to leading criminology and criminal justice scholars knowledgeable in the area of broken windows policing strategies. This will help to identify studies the above searches


left out as these experts may be able to make referrals to studies that were missed, particularly unpublished studies.

Finally, an information specialist will be engaged at the outset of our review and at points along the way in order to ensure that appropriate search strategies were used to identify the studies meeting the criteria of this review.\(^3\) For instance, we will work with the information specialist to conduct an extensive Google search for eligible studies that will seek to identify eligible studies by using the search terms below as well as including words such as “research,” “evaluation,” and “program analysis.” The information specialist will be consulted on the use of Google Scholar to identify studies that cite seminal broken windows studies (e.g. Wilson & Kelling, 1982; Kelling & Coles, 1996) and on the use of publisher databases and indexes (e.g. Wiley, Sage, and Springer) to identify potentially eligible studies.

The following fifteen databases will be searched:

1. Criminal Justice Periodical Index
2. Sociological Abstracts
3. Social Science Abstracts (SocialSciAbs)
4. Social Science Citation Index
5. Arts and Humanities Search (AHSearch)
6. Criminal Justice Abstracts
8. Educational Resources Information Clearinghouse (ERIC)
9. Legal Resource Index
10. Dissertation Abstracts
12. Google Scholar
13. Online Computer Library Center (OCLC) SearchFirst
14. CINCH data search
15. Academic Search Premier

The following terms will be used to search the fifteen databases listed above:

1. Broken windows AND police
2. Disorder AND police
3. Incivilities AND police
4. Disorder policing
5. Order maintenance policing
6. Zero tolerance policing
7. Quality of life policing
8. Misdemeanor arrest policing
9. Signal crimes

---

\(^3\) Ms. Phyllis Schultze of the Gottfredson Library at the Rutgers University School of Criminal Justice will be consulted on the initial abstract search and contacted throughout on our search strategies.
In addition, two existing registers of randomized controlled trials will be consulted. These include (1) the “Registry of Experiments in Criminal Sanctions, 1950-1983 (Weisburd et al., 1990) and (2) the “Social, Psychological, Educational, and Criminological Trials Register” or C2 SPECTR being developed by the United Kingdom Cochrane Centre and the University of Pennsylvania (Turner et al., 2003). Two additional online databases of rigorous studies in policing will be reviewed: the Evidence-Based Policing Matrix (http://www.policingmatrix.org) and the US Office of Justice Program’s CrimeSolutions.gov website.

Selection of studies

The reviewers will screen abstracts and leads to potentially eligible studies and decide which full-text reports should be acquired. Only the full-text papers of titles and abstracts indicating, or potentially indicating, an evaluation of a broken windows policing or an empirical analysis of the theoretical connections between disorder and crime will be obtained. In cases of ambiguity, the full text of the study will be obtained in order to properly determine whether an eligible study design was used. Studies that use randomized controlled or rigorous quasi-experimental designs will be considered for inclusion in the main review. Correlation and observational studies without control groups will be noted and analyzed separately. These descriptive studies will not be included in the formal analysis reporting the findings of the review.

Assessment of Methodological Quality and Bias

Studies meeting the eligibility set forth above will be coded for a range of characteristics related to methodological quality including the definition criteria used to identify the units of analysis, the statistical tests used to determine crime prevention effectiveness, the measurement of displacement, the violation of randomization procedures, case attrition from the study, and the subversion of the experiment by participants. Farrington (2003) proposes five easily understood methodological criteria to assess the methodological quality of evaluation studies. These criteria include statistical conclusion validity, internal validity, construct validity, external validity, and descriptive validity.

As appropriate and possible, the role of the various methodological factors on the observed empirical results will be assessed. The methodological quality information will be reported in tabular form and presented along with the effect size information. If a meta-analysis is performed, effect sizes will be compared on the above dimensions to assess for potential bias. However, it is important to recognize that eligible studies may not detail or even mention implementation issues. Indeed, all field experiments face implementation difficulties and care will be taken not to artificially downgrade the value of certain studies simply because one provided an open account of potential process problems.
Publication bias presents a strong challenge to any review of evaluation studies (Rothstein, 2008). Campbell reviews, such as ours, take a number of steps to reduce publication bias. Our extensive search procedures, the use of an information retrieval specialist (Phyllis Schultze), and the mobilization of an extensive network of police scholars make it unlikely that relevant unpublished works would remain hidden from this review. Nevertheless, we will use the trim-and-fill procedure (Duval & Tweedie, 2000) to estimate the effect of potential data censoring, such as publication bias, on the outcome of the meta-analyses.

Data Management and Extraction

The reviewers, with the help of a trained research assistant, will extract information from the full text report on the characteristics of the study using a carefully designed data extraction instrument (see coding instruments included in Appendix A). A content analysis will be conducted on the full text of the report and the data extraction instrument will capture data on the relevant dimensions of this review. These dimensions include: a complete description of the treatment, methods used to define and identify targeted areas, research design and statistical techniques, threats to the research design, crime outcome measures, and alternative outcome measures. When important information is missing from available study reports, the original researchers will be contacted, if possible, to determine if they can supply that information.

The two review authors, Braga and Welsh, will independently code each eligible study. The authors will then compare the two completed coding instruments for each eligible study to identify any divergent coding of study material. Where there are discrepancies, the authors will jointly review the study and determine a final coding decision. This process will identify and resolve any coding differences.

Finally, pre-test information will be coded for eligible quasi-experimental studies when available. This will allow the calculation of pre-post effect sizes for eventual inclusion in the meta-analyses described below.

Determination of Independent Findings

A single evaluation of broken windows policing intervention may provide data on multiple outcome measures. For example, the randomized controlled trial of the Lowell policing disorder intervention presents an array of outcome measures including assault, robbery, burglary / breaking and entering, larceny / theft, disorder / nuisance, and total citizen calls for service data (Braga & Bond, 2008). A separate study reported on the community perceptions of the value of the policing disorder intervention in addressing neighborhood crime and safety conditions (Braga & Bond, 2009). Policing disorder interventions are often targeted at underlying environmental conditions that may be responsible for generating a variety of offenses such as violent crimes, property crimes, and drug crimes. The treatment could have varying effects on offending trends and patterns in different crime categories.
For cases such as this with multiple findings from the same sample, each will be examined independently to decide how to either combine the findings or to choose the one that best represents the study. Some policing disorder interventions may be designed to deal with a specific problem, but other programs may also target some secondary problems and report outcomes for these as well. In these cases the effect size for the primary problem will be reported. For instance, while multiple outcomes are reported, the Lowell evaluation clearly identifies the pre-test and post-test counts of total citizen calls for service in the treatment hot spots relative to the control hot spots as the primary outcome measure (Braga & Bond, 2008).

However, some studies might have multiple primary outcomes. Analyzing these separately would clearly lead to problems regarding statistical dependence of outcomes. As such, we will code a maximum of three primary outcomes, with the criteria of choosing the maximum, moderate/median effect size and minimum effect size to offer flexibility in calculating an overall effect size for such studies. The same strategy will be used for any studies reporting the same outcome multiple times with different types of data (i.e. a study evaluating the impact of a policing disorder on “robbery” may use robbery incidents and robbery calls for service as primary outcome measures). Finally, some studies may involve multiple sites—i.e. policing disorder program delivered by one police department/taskforce to specific problems in multiple areas within a city. Such cases will be treated as one study with sub units, and independent effect sizes for primary outcomes will again be created in the same manner as above.

**Statistical Procedures and Conventions**

Analysis of outcome measures across studies will be carried out in a uniform manner and, where appropriate and possible, involve quantitative analytical methods. Meta-analytic procedures will be used to combine data from studies. We will be limiting our meta-analyses to studies that evaluate the impacts of broken windows policing on crime outcomes at the area level. As described by Lipsey and Wilson (2001), it is problematic to combine effect sizes from studies with very different units of analysis (such as combining studies focused on people with studies focused on places).

For eligible studies, with enough data present, effect sizes will be calculated using the standardized measures of effect sizes as suggested in the meta-analytic literature (e.g. see Lipsey & Wilson, 2001). Mean effect sizes will be computed across studies and we will use a correction such as the inverse variance weight for computing the associated standard error. The odds-ratio will be the effect size of choice for all outcomes of a dichotomous or binary nature. For eligible studies that contrast two groups that have a continuous underlying distribution, the standardized mean difference effect size (also known as Cohen’s $d$; see Rosenthal, 1994) will be used in our meta-analysis to synthesize results. This effect size measure would be appropriate for eligible evaluations that contrast an outcome measure, such as mean incident counts, for treatment and control areas over pre-test and post-test observation periods. The Effect Size Calculator,
developed by David B. Wilson and available on the Campbell Collaboration’s web site, will be used to calculate odds-ratio and standardized mean difference effect sizes for reported outcomes in each study. Biostat’s Comprehensive Meta Analysis Version 2.2 will then be used to conduct the meta-analysis of effect sizes. In our proposed meta-analyses, we will use a random-effects model.

We will examine four contextual or moderating features of broken windows policing approaches. To assess these contextual or moderating features, we will use the analog to the ANOVA method of moderator analysis (see Lipsey & Wilson, 2001) for categorical moderator variables and meta-analytic regression analysis for continuous moderator variables or analyses involving multiple moderators.

First, consistent with other systematic reviews of policing interventions (e.g. Braga & Weisburd, 2012), we will examine research design type as a moderator of program effect sizes. Relative to quasi-experimental designs, randomized controlled trials have been revealed to generate more conservative effect sizes (Weisburd et al., 2001; Welsh et al., 2011). Second, we will examine policing disorder program type as a moderator of program effect sizes. As our literature review suggests, there seems to be varying approaches that police departments take to controlling disorder that could produce different effects on crime (e.g., zero-tolerance misdemeanor arrest strategies v. community problem-solving approaches to recurring disorder problems). Third, we will conduct a moderator analysis of the impacts of policing disorder programs on different types of crime outcomes (e.g. violent crimes v. property crimes). It is possible that a police focus on disorder could produce disparate impacts on particular types of crimes.

Fourth, and finally, we will examine the unit of analysis as a moderator of program effect sizes. It is possible that the types of areas (e.g. small crime places vs. larger neighborhoods) subjected to policing disorder strategies will matter in program effect sizes. Consistent with the literature on hot spots policing (Braga, 2007), broken windows interventions concentrated in small areas may generate stronger crime reduction impacts relative to treatments diffused across larger geographic units. Therefore, we will conduct separate analyses based on the geographical unit (e.g. city-wide, neighborhood, precinct, or defined hot spot place such as street corners and block faces).

Treatment of Qualitative Research

Qualitative research on crime and disorder outcomes will not be formally included in this systematic review. Qualitative insights on the crime prevention value of broken windows policing will be included as descriptive information in the review report. If the search

---

4 http://www.campbellcollaboration.org/resources/effect_size_input.php

5 We will examine the Q statistic to assess heterogeneity of effect sizes across studies. It is possible that a large population of policing disorder program studies may produce a distribution of crime control effects are fairly consistent one to another. However, our brief literature also suggests that there may be considerable diversity across these programs. Fortunately, the random effects model converges on the fixed-effects model as the distribution becomes homogeneous.
strategies reveal a number of qualitative studies, the authors will engage a qualitative researcher to assist in future updates to this review with a synthesis of qualitative evaluation measures.

**Timeframe**

The estimated timeline for a completed report includes the following benchmarks and anticipated dates:

<table>
<thead>
<tr>
<th>Task</th>
<th>Date</th>
</tr>
</thead>
<tbody>
<tr>
<td>Search for published and unpublished studies</td>
<td>June 2012</td>
</tr>
<tr>
<td>Relevance assessments</td>
<td>July 2012</td>
</tr>
<tr>
<td>Coding studies and resolving coding differences</td>
<td>August 2012</td>
</tr>
<tr>
<td>Extraction of data from research reports</td>
<td>August 2012</td>
</tr>
<tr>
<td>Statistical analysis</td>
<td>October 2012</td>
</tr>
<tr>
<td>Preparation of report</td>
<td>November 2012 – January 2013</td>
</tr>
<tr>
<td>Submission of completed report</td>
<td>February 2013</td>
</tr>
</tbody>
</table>

**Updating the Review**

In accordance with Campbell Collaboration guidelines, we will update this review once every three years.

**Conflict of Interest**

With colleagues, Braga has conducted two randomized controlled trials that found disorder policing, implemented within a problem-oriented policing framework, to be effective in reducing citizen calls for service to the police in crime hot spots (see Braga et al., 1999; Braga & Bond, 2008). Although Braga doesn’t have an ideological bias towards the effectiveness of broken windows policing interventions, it may be uncomfortable for him to report findings in this review that contradict the findings of his evaluations or related evaluations conducted by his colleagues. Welsh does not have any conflicts of interest in conducting this review.
References


Executive Research Forum.


Karmen, A. (2000). New York murder mystery: The true story behind the crime crash of


APPENDIX A. CODING PROTOCOL

Reference Information

1. Document ID: __ __ __ __

2. Study author(s):

3. Study title:

4. Publication type: ______
   1. Book
   2. Book chapter
   3. Journal article (peer reviewed)
   4. Thesis or doctoral dissertation
   5. Government report (state/local)
   6. Government report (federal)
   7. Police department report
   8. Technical report
   9. Conference paper
   10. Other (specify))________________________

5. Publication date (year): ____________

6a. Journal Name:

6b. Journal Volume: ________________

6c. Journal Issue: ________________

7. Date range of research (when research was conducted):
   Start: ______________
   Finish: ______________

8. Source of funding for study:

9. Country of publication: ________________

10. Date coded: ______________

11. Coder’s Initials: __ __ __
Describing the Broken Windows Policing Intervention

12a. Did the study formally identify the treatment as a broken windows policing intervention?
   1. Yes
   2. No

12b. If No, what did the study call the intervention?
   __________________________________________

13. What crime problem was targeted for the intervention? (Select all that apply)
   1. Total crime
   2. Violent crime
   3. Property crime
   4. Homicide
   5. Sexual assault / rape
   6. Robbery
   7. Assault
   8. Burglary
   9. Larceny
   10. Motor vehicle theft
   11. Disorder
   12. Other (specify) __________

14. What unit of analysis was used? (Select all that apply)
   1. Crime hot spots (specific small places)
   2. Police-defined geographical unit (districts, precincts, etc.)
   3. U.S. Census-defined geographical unit (tracts, block groups, etc.)
   4. Neighborhoods as defined by researchers
   5. Other unit (specify) __________

15. What type of broken windows policing intervention was implemented in the targeted areas? (Select all that apply)
   1. Community input on disorder problems
   2. Increased misdemeanor arrests to control social disorder
   3. Situational strategies to modify physical environment
   4. Social service strategies to deal with disorderly persons (substance abusers, homeless, mentally-ill offenders, etc.)
   5. Other (specify) __________
16. What did the evaluation indicate about the implementation of the response? ______
   1. The response was implemented as planned or nearly so
   2. The response was not implemented or implemented in a radically different way than originally planned
   3. Unclear/no process evaluation included

17. If the process evaluation indicated there were problems with implementation of the response, describe these problems:
________________________________________________________________________
________________________________________________________________________
________________________________________________________________________

18. Country where study was conducted: ________________________________

19. City (and state/province, if applicable) where study was conducted:
________________________________________________________________________
________________________________________________________________________

Methodology/Research design:

20. Type of study: __________
   1. Randomized experiment
   2. Nonequivalent control group (quasi-experimental)
   3. Multiple time series (quasi-experimental)
   4. Other (specify) ___________________

21. How were study units allocated to treatment or comparison conditions?
   1. Simple random allocation
   2. Random allocation in pairs, blocks, or some other sophisticated technique
   3. Simple descriptive matching
   4. Sophisticated statistical matching (e.g. propensity scores)
   5. Other (specify) ___________________

22. Explain how independent and extraneous variables were controlled so that it was possible to disentangle the impact of the intervention or how threats to internal validity were ruled out.
________________________________________________________________________
23. Did the study measure spatial crime displacement and diffusion of crime control benefits?
   1. Yes
   2. No

24. Explain how the study measured spatial crime displacement and diffusion of crime control benefits.

The following questions refer to the units receiving treatment:

25. Units receiving treatment: ______
   1. Crime hot spots (specific small places)
   2. Police-defined geographical unit (districts, precincts, etc.)
   3. U.S. Census-defined geographical unit (tracts, block groups, etc.)
   4. Neighborhoods as defined by researchers
   5. Other unit (specify) __________

26. What is the exact unit receiving treatment?

The following question refers to the units not receiving treatment

27. Units NOT receiving treatment: ______
   1. Crime hot spots (specific small places)
   2. Police-defined geographical unit (districts, precincts, etc.)
   3. U.S. Census-defined geographical unit (tracts, block groups, etc.)
   4. Neighborhoods as defined by researchers
   5. Other unit (specify) __________

28. What were the casual hypotheses tested in this study?
29. Please identify any theories from which the causal hypotheses were derived.

________________________________________________________________________
________________________________________________________________________
________________________________________________________________________

Outcomes reported (Note that for each outcome, a separate coding sheet is required. This includes main effects outcomes as well as crime displacement and diffusion of crime control benefits outcomes)

30. How many crime / alternative outcomes are reported in the study? ______

31. What is the specific outcome recorded on this coding sheet?
________________________________________________________________________

32. Was it the primary outcome of the study? ______
   1. Yes
   2. No
   3. Can’t tell/researcher did not prioritize outcomes

33. Was this initially intended as an outcome of the study? ______
   1. Yes
   2. No (explain)
   3. Can’t tell

34. If no, explain why:
________________________________________________________________________

Unit of analysis

35. What was the unit of analysis for the research evaluation?
   1. Crime hot spots (specific small places)
   2. Police-defined geographical unit (districts, precincts, etc.)
   3. U.S. Census-defined geographical unit (tracts, block groups, etc.)
   4. Neighborhoods as defined by researchers
   5. Other unit (specify) ___________

36. How many units of analysis are there for the intervention in the study? ______

37. Did the researchers collect nested data within the unit of analysis?
   1. Yes
   2. No
**Dependent Variable**

38. What type of data was used to measure the outcome covered on this coding sheet?
   1. Official data (from the police)
   2. Researcher observations
   3. Self-report surveys
   4. Other (specify) ___________________

39. If official data was used, what specific type(s) of data were used? (Select all that apply)
   1. Calls for service (911 calls)/crime reports
   2. Arrests
   3. Incident reports
   4. Level of citizen complaints
   5. Other (specify)
   6. N/A (official data not used)
   7. Other (specify) ___________________

40. If researcher observations were used, what types of observations were taken? (Select all that apply)
   1. Physical observations (e.g. observed urban blight, such as trash, graffiti)
   2. Social observations (e.g. observed disorder, such as loitering, public drinking)
   3. Other observations (specify)
   4. N/A (researcher observations not used)
   5. Other (specify) ___________________

41. If self-report surveys were used, who was surveyed? (Select all that apply)
   1. Residents/community members
   2. Business owners
   3. Elected officials
   4. Government/social service agencies
   5. Other (specify) ___________________
   6. N/A (self-report surveys not used)

42. Specifically identify the outcome covered on this coding sheet

______________________

43. For the units of analysis in this study, what time periods were examined for the outcome covered on this coding sheet?
   1. Yearly
   2. Monthly
   3. Weekly
   4. Other researcher defined time periods (specify)

______________________
44. What was the length in time of the follow-up period after the intervention?

___________________________________________

45. Did the researcher assess the quality of the data collected?
   1. Yes
   2. No

46a. Did the researcher(s) express any concerns over the quality of the data?
   1. Yes
   2. No

46b. If yes, explain

___________________________________________

___________________________________________

Effect size/Reports of statistical significance

Dependent Measure Descriptors

47. Statistical analysis design: _____
   1. Pretest comparison
   2. Post-test comparison
   3. Follow-up comparison
   4. N/A

Sample Size

48. Based on the unit of analysis for this outcome, what is the total sample size in the analysis?

_______

49. What is the total sample size of the treatment group (group that receives the response)?

_______

50. What is the total sample size of the control group (if applicable)? _____

51a. Was attrition a problem in the analysis for this outcome?
   1. Yes
   2. No
51b. If attrition was a problem, provide details (e.g. how many cases were lost and why were they lost).

_________________________________________________________________________
_________________________________________________________________________

52. What do the sample sizes above refer to?
   1. Crime hot spots (specific small places)
   2. Police-defined geographical unit (districts, precincts, etc.)
   3. U.S. Census-defined geographical unit (tracts, block groups, etc.)
   4. Neighborhoods as defined by researchers
   5. Other unit (specify) __________

Effect Size Data

53. Raw difference favors (i.e. shows more success for):
   1. Treatment group
   2. Control group
   3. Neither (exactly equal)
   4. Cannot tell (or statistically insignificant report only)

54. Did a test of statistical significance indicate statistically significant differences between either the control and treatment groups or the pre and post tested treatment group? _____
   1. Yes
   2. No
   3. Can’t tell
   4. N/A (no testing completed)

55. Was a standardized effect size reported?
   1. Yes
   2. No

56. If yes, what was the effect size? ______

57. If yes, page number where effect size data is found ________

58. If no, is there data available to calculate an effect size?
   1. Yes
   2. No

59. Type of data effect size can be calculated from:
   1. Means and standard deviations
   2. t-value or F-value
   3. Chi-square (df=1)
4. Frequencies or proportions (dichotomous)
5. Frequencies or proportions (polychotomous)
6. Other (specify) _________

Means and Standard Deviations

60a. Treatment group mean. _____
60b. Control group mean. _____

61a. Treatment group standard deviation. _____
61b. Control group standard deviation. _____

Proportions or frequencies

62a. \( n \) of treatment group with a successful outcome. _____ 
62b. \( n \) of control group with a successful outcome. _____ 

63a. Proportion of treatment group with a successful outcome. _____
63b. Proportion of treatment group with a successful outcome. _____

Significance Tests

64a. \( t \)-value _____
64b. \( F \)-value _____
64c. Chi-square value \((df=1)\) _____

Calculated Effect Size

65a. Effect size ______
65b. Standard error of effect size _____

Conclusions made by the author(s)

Note that the following questions refer to conclusions about the effectiveness of the intervention in regards to the current outcome being addressed on this coding sheet.

66. Conclusion about the impact of the broken windows intervention? _____
   1. The authors conclude the program positively impacted crime / disorder
   2. The authors conclude the problem did not positively impact crime / disorder
   3. Unclear/no conclusion stated by authors

67. Did the assessment find evidence of a geographic displacement of crime? ______
   1. Yes
2. No
3. Not tested

68. Did the assessment find evidence of other types of displacement of crime? _____
   1. Yes. Please specify ______________
   2. No
   3. Not tested

69. Did the assessment find evidence of a geographic diffusion of crime control benefits? _______
   1. Yes
   2. No
   3. Not tested

70. Did the assessment find evidence of other types of diffusion of crime control benefits? _____
    1. Yes. Please specify ______________
    2. No
    3. Not tested

71. Did the author(s) conclude that the broken windows policing intervention was beneficial? __
    1. Yes
    2. No
    3. Can’t tell

72. Did the author(s) conclude there was a relationship between the broken windows policing intervention and a reduction in crime? ______
    1. Yes
    2. No
    3. Can’t tell

73. Who funded the intervention?
   __________________________________________________________
   __________________________________________________________

74. Who funded the evaluation research?
   __________________________________________________________
   __________________________________________________________

75a. Were the researchers independent evaluators?
    1. Yes
    2. No
75b. If no, explain the nature of the relationship:


76. Additional notes about conclusions:


77. Additional notes about study: