EDITORIAL INTRODUCTION

GANG DATABASES

Gang databases

To be or not to be

Irving Spergel, *Senior Editor*
The University of Chicago

J ulie Barrows and C. Ronald Huff's (2009, this issue) review of state-gang statutes and
their recommendation of the Minnesota model favor the use of common criteria to iden-
tify gangs, categories of gang members, and gang database procedures. James B. Jacobs,
David M. Kennedy, Malcolm W. Klein, and James W. Short, Jr., in each responding policy
essay (2009, this issue, respectively), raise strong questions about validity, reliability, as well
as public-policy utility of gang law-enforcement definitions and gang databases. The review
and responses are contributions to the current vigorous debate regarding proposed legislation
in the U.S. House of Representatives and the U.S. Senate that is supported and opposed by
several interested local, state, and national governmental and nongovernmental organizations.
Underlying concerns are how best to address the present and expected-to-be-growing gang
problem in the United States.

In my editorial remarks, I indicate the nature of the general oppositional approaches to
the gang problem in the proposed House and Senate legislation. Next, I draw attention to the
contribution of the Barrows–Huff (2009) statement but raise questions about the recommended
definitions of a gang member in their Minnesota model. Then, I emphasize what I consider
to be key points in the four responding policy essays. Finally, I introduce more observations
and several findings of my in-progress analysis of a cross-site, integrated six-city evaluation of
a comprehensive, community-wide gang program approach to reducing the gang problem,[1]
which bears on gang-member definitions and how to address the gang problem.

Direct correspondence to Irving Spergel, School of Social Service Administration, The University of Chicago, 969
East 60th Street, Chicago, IL 60637 (e-mail: iasperge@uchicago.edu).

1.   Kwai Ming Wa and I are integrating and analyzing data collected from and about approximately 1,600
      gang youth in six programs (Bloomington-Normal, IL, Mesa and Tucson, AZ, San Antonio, TX, and Riverside,
      CA) that tested the Office of Juvenile Justice and Delinquency Prevention's (OJJDP's) Comprehensive,
      Community-Wide Approach to Gang Prevention, Intervention and Suppression initiative (1995–2000),
      and the earlier prototype program in Little Village, Chicago, IL, funded by the U.S. Department of Justice's
      Violence in Urban Areas Program through the Illinois Criminal Justice Information Authority (1992–1997).

Senate and House Dissensus

The Gang Abatement and Prevention Act of 2007 (S. 456, 110th Cong.) was passed in September 2007, but its equivalent in the U.S. House of Representatives was not. The same Senate Bill (now S.132, 111th Cong.) was reintroduced in 2009, as was an equivalent House Bill (Gang Prevention, Intervention and Suppression Act of 2009, H.R. 1022, 111th Cong.). At this writing, both bills have not been enacted. The Senate Bill (S.132) is based on extensive local, state, and federal information provided and supported by law-enforcement agencies and mayors' offices. It reports that gangs are present in every state, that gangs account for a substantial proportion of homicides, and that some gangs recently are aligned with or working for Mexican violent drug gangs and comprise a growing threat to law and order. S. 132 proposes new federal laws and enhanced penalties and sentences for members of identified criminal gangs that have committed and were convicted of serious enumerated crimes. Federal and local agencies will be funded to collaborate to address the problem in High Intensity Gang Areas across states. Less funding would be provided to local community-based gang prevention and intervention programs. A center for gang research and evaluation also would be established.

The U.S. Government Accountability Office (GAO) report (2009), *Combating gangs*, directs more specific attention to the need for better coordination of gang data systems as well as performance measurement between the Department of Justice and the Department of Homeland Security. At least 15 units of these two organizations currently are engaged in collecting and identifying data and organizing gang data systems without adequate coordination within and across these departments.

A different and unrelated analysis and proposal for action was presented in the House of Representatives through the Youth Prison Reduction through Opportunities, Mentoring, Intervention, Support, and Education or the Youth Promise Act of 2007 (H.R. 3846, 110th Cong.) but not passed. It has been re-introduced (H.R. 1064, 111th Cong.) in 2009. H.R. 1064 does not focus on the adult criminally organized gang problem, as does S. 132 (2009), but on a broad array of youth delinquency and youth gang problems. The House Bill (H.R. 1064) focuses broadly on youth social development, which includes prevention, education, gang intervention, mentoring, job training, health and mental-health services, and community- and school-based programs. It opposes the "standard suppression and incapacitation" approach of law enforcement. It would establish a Center for Youth-Oriented Policing as well as a national center for juvenile justice and delinquency prevention proven practices research. H.R. 1064 recommends that several federal agencies, which include the following departments—justice, education, labor, housing and urban development, defense, health and human services, and so on—be involved in addressing the youth-gang and youth-delinquency problems. The OJJDP would guide the development of prevention and intervention programs.

Each of these approaches claims to be comprehensive, although S. 132 (2009) would emphasize the development of interagency councils of mainly law enforcement agencies at national and interstate levels; H.R. 1064 (2009) would emphasize comprehensive approaches that involve a broad array of federal agencies and the use of local interagency councils to focus

on the coordination of community-based programs. The Senate Bill (S. 132) would require development of a common database focused on gathering intelligence and on identifying current gangs and threat groups as well as hard-core gang leaders and gang influentials. It would not accord major attention to youth or community-development needs—and the changes necessary—as components of the control and reduction of the gang problem. H.R. 1064 would focus on the development of decentralized systems of youth databases that identify the needs of children and adolescent youth at risk and provide for a range of social services and for appropriate educational, job, and health opportunities.

The problems and interests of children and youth as a component of the gang problem would generally be ignored in S. 132 (2009), but the role of criminal gangs and adults would be ignored in H.R. 1064 (2009). Most individuals listed in gang databases by local law enforcement are probably middle- and old-age adolescents, at least at initial contact. The nature of the interrelationship of juveniles, youth, and adults in the development of criminal gangs is not addressed clearly in either approach. The concepts of delinquents and juvenile gangs versus organized criminal gangs have been abstracted and located—for policy and legislative purposes—in separate gang worlds. In reality, the supposedly different gang problems are variably and complexly interrelated. Gang policy and practice have to be oriented and implemented in some coordinated way to achieve social development and treatment as well as suppression and incapacitation. Social-development administrators and practitioners, especially law-enforcement policymakers, have a major responsibility to reach out and facilitate coordination and collaboration of contacts and services for the youth-majority identified in gang databases.

Barrows–Huff Statement

The Barrows–Huff (2009) review of the gang literature, analysis of state-level gang databases, and recommendation of the Minnesota-model gang statute generally is consistent with the framework of S. 132 (2009), except for its greater emphasis on local and state rather than on national definitions of the gang problem. They, and other gang experts, observe that there have been many efforts to develop a clear and common definition of a gang and a gang member in state statutes, to be implemented by law enforcement departments. Barrows and Huff indicate, nevertheless, that few states have statutes "governing the compilation, maintenance, and dissemination of gang information." They see a need to improve and enhance state definitions and, especially, the construction of gang databases at the state level, based on the Minnesota Model. On the one hand, they see problems in the implementation of a federal definition and a gang database by federal agencies to apply across current states and localities, which are so different. On the other hand, they urge the creation of a federal commission composed of local, state, and national officials and experts to consider the issues involved.

The policy essays by Jacobs, Kennedy, Klein, and Short (2009, respectively) raise questions about the concept and use of law enforcement gang databases. They question, as I do specifically, the criteria used in the Minnesota Model's three categories of gang member: "suspected gang member," "confirmed gang member," and "confirmed and convicted gang member." The

definitions are applied based on ten criteria, listed separately and in combination, although only self-admission has been regarded by researchers as the "best measure" of gang member-ship (Esbensen, Winfree, He, and Taylor, 2004). Other criteria of gang member cited in the Minnesota Model include "associate[s] on a regular basis with known gang members" and "cor-responds with known gang members," which could lead to classification of a mother or other relative or a girlfriend or interested neighbor, teacher, or priest as a gang member. The label of gang member could remain with the individual throughout the years, even after the individual listed as a "confirmed gang member" or "confirmed and convicted gang member" is no longer associated with a gang, or regardless of whether later he or she commits a crime or commits a crime that is not gang related.

Perhaps Barrows–Huff (2009), in their reference to the federal "Violent Gang and Terrorist Organization File (VGTOF)," and as indicated in California and other state legislation, should have addressed the issue of the relation of the gang to terrorism. As yet, no substantial evidence has been found that U.S. gangs intend to attack or overthrow local, state, or federal government and their institutions. Gang members often disregard, violate, and manipulate laws, but they do not directly or expressly seek to replace them. A research report from the Washington Of-fice on Latin America (WOLA) (Gray, 2009) stated that "Sensational accounts by the media, politicians, and police have exaggerated the connection between Central American youth gangs and drug trafficking, international organized crime, and terrorism."

The Challenge to Gang Criteria and Databases

The policy essays by Jacobs, Kennedy, Klein, and Short (2009, respectively) are concerned about the validity, reliability, and feasibility of gang definitions and construction of gang data-bases. Jacobs (2009) raises civil liberty questions. He states that "police, jails, and prisons can label an individual as a gang member without affording him or her any opportunity to contest the label and without notifying the individual that the labeling has occurred." He wonders whether gang-related information can ever be removed from a gang database: "[I]t should not be surprising to find a large discrepancy between how the gang database should operate and how it does operate."

Kennedy (2009) says that Barrows and Huff (2009) "get it right when it comes to the prob-lems with" gang databases. He says that gang databases may be "overinclusive, underinclusive, poorly conceived in particular places, inconsistent from place to place, guided by statutes and policies … and often implemented in practice in ways that are inconsistent with these statutes and policies." Kennedy finds in his research that groups and individuals not classified as gangs or gang members may be at "the heart of serious violent crime, drug activity, and a wide range of other core public-safety problems." He concludes: "It is not clear, however, that we need to solve the gang definition problem and create coherent, far less nationally consistent, gang databases to solve them." He advises that it may be more useful to ignore the gang designation and focus on the nature of the criminal act, regardless of the gang label of who commits it.

Klein (2009) states that gang databases at the state level are aggregations of local police-run databases. The local databases "suffer seriously from problems of reliability and validity, along with major variations in content." He notes that, among other deficiencies of gang databases is the underreporting of female members and crime as well as the missing of short-tenure gang members and inconsistent gang members. He observes that younger youth-gang members are less likely than older youth-gang members to be visible to the police. Klein reports that "U.S. cities are replete with both public and private programs to prevent, intervene in, or suppress gang activity. Most programs are not evaluated for effectiveness." He refers to a "massive new gang prevention and intervention program …being initiated by the Los Angeles Mayor's Office." He speaks of research questions, as yet unresolved, such as outcome measures using police data versus self-reported data, or both, and "how to meld the two."

Short (2009) states that databases should be inclusive, broadly based on reports not only by police or other law-enforcement agencies but also by "databases that involve other institutions and other levels of expertise." Referring to Sullivan's (2006) work on reification, Short indicates that "[r]eification of gangs occurs at every level: among local community residents and the general public, legal institutions, academic research, and gang members." He warns against concentrating on the outward symbols of gang membership without addressing "local community contexts." He cautions against "direct social policy application of any gang definition or identification without careful, systematic attention to local community contexts."

Additional Comments

The gang problem, especially the youth component of the problem in various forms, has become increasingly widespread and serious in the United States and in many developed and developing countries of the world. We have accumulated large amounts of information about the nature and scope of the problem, based on media, legislative reports, and observations and surveys by researchers using police data and short-term case studies. Similar conditions and pressures seem to affect the development of somewhat varied gang problems inside and outside of the United States. Much more knowledge and research is required related to policy, program, and practice issues. My view is that the gang problem is a function not only of the interaction of population movement, low income, low status, minority race/ethnic, community and family patterns of social disorganization, and political opportunism, but also it is a function of youth (mainly male) social-development factors—all interacting with each other.

Many of these interacting factors must be taken into account in the definition of a gang, an individual gang member, a gang incident, and the construction and use of gang databases. Gang definitions and databases serve not only theoretical and academic purposes but also, primarily and more often, the policy, personal, and political purposes of groups, organizations, communities, and legislatures that are often not congruent with each other. The means to improving definitions and gang databases, therefore, lie not only in developing knowledge about the behavior of gangs and their contexts—economic, social, political, and cultural—including the processes of youth development but also in how best to create consensus and improve

coordination about definitions, construct and use gang databases, and address the problem in ways sensitive to the framework of its complex dimensions.

We also need elaborate definitions of gang, gang member, and databases in terms of how to address their relationships to the problem of social context. One example might be the expansion of the Eurogang definition of a youth gang, which was recently created by a consortium of U.S. and European gang experts (Van Gemert, Peterson, and Lien, 2008). Their definition could be elaborated as follows: A youth gang is any durable and variably organized street-oriented group of juveniles, mainly adolescents, and young adults, located in, perceived to be, and deriving from socioeconomically and culturally marginalized community and institutional contexts, which may include the presence and influence of adult criminal organization.

There are many youth age-category influences on the development of the gang problem. Younger kids in gangs, if they sustain gang membership, could become violent and seriously criminally involved; older youth generally age out of gang behavior. The possibility that connections of gang youth to neighborhood adults and their own families could be related to criminal organizations should not be overlooked. Furthermore, although gangs, individual gang members, and the criminal incidents they commit might differ somewhat across sites, categories of types of gangs, gang members, and criminal incidents can still be constructed and used across sites. For example, some gang youth who are gang participants may not have self-reported offense or criminal arrest records; those from the same gang who have arrest records may have been relatively more engaged in either violent, property or drug-selling crimes, or combinations of these behaviors (Spergel, 2007).

Special gang databases should be constructed, with the aid of research and policy personnel, to test and demonstrate how to create and focus attention on the reduction of the gang problem. Federal guidance and regulation from both the Department of Justice and the Department of Health and Human Services should guide and support this process, along with close monitoring. A variety of interest groups at national, state, and local levels will be involved in an advisory capacity. They will be concerned variously with community protection, deterrence, incapacitation of gang behavior, as well as social development, prevention, intervention, and rehabilitation of gang members. Hopefully, through the processes of interaction of these interest groups, appropriate collaboration of strategies will begin to occur and be sustained at interagency, program-structure, and worker-practice levels.

In our previous comprehensive gang-program evaluation research (Spergel, 2007; Spergel, Wa, and Sosa, 2006),[2] and current reanalysis, we have made use of integrated data sets derived from detailed youth interview, program worker, and police arrest records. Program and comparison youth were matched across sites. Propensity scoring in addition to general linear and linear mixed models have been used to analyze data. It is possible that such a research-developed gang database model, sensitive to interacting suppression and social-development approaches, can be modified for potential interactive use by law enforcement and social-development interests at policy and practice levels, under adequate restrictions and regulations. In any case, the values of gang databases should be tested through more complex, long-term program policy and research collaboration than exists at the present time.

References

Barrows, Julie C. and C. Ronald Huff. 2009. Gangs and public policy: Constructing and deconstructing gang databases. *Criminology & Public Policy*. This issue.

Esbensen, Finn-Aage, L. Thomas Winfree, Jr., Ni He, and Terrence J. Taylor. 2004. Youth gangs and definitional issues: When is a gang a gang, and why does it matter? In (Finn-Aage Esbensen, Stephen G. Tibbetts, and Larry Gaines, eds.), *American youth gangs at the millennium*. Long Grove, IL: Waveland.

Gray, Guthrie. 2009. *Politics: Crackdown on Latin youth gangs misdirected*. Washington, DC: Washington Office on Latin America.

Jacobs, James B. 2009. Gang databases: Context and questions. *Criminology & Public Policy*. This issue.

Kennedy, David M. 2009. Comment on "Gangs and public policy: Constructing and deconstructing gang databases". *Criminology & Public Policy*. This issue.

Klein, Malcom W. 2009. Street gang databases: A view from the gang capitol of the United States. *Criminology & Public Policy*. This issue.

Short, James F., Jr. 2009. Gangs, law enforcement, and the academy. *Criminology & Public Policy*. This issue.

Spergel, Irving A. 2007. *Reducing youth gang violence: The Little Village Gang Project in Chicago*. Lanham, MD: Altamira.

2. In the integrated data analysis of the effects and effectiveness of a comprehensive program approach across six sites, approximately 800 program youth and approximately 800 non-served comparison youth were matched individually. Data were derived using variables from interviews, self-reports, and police arrest histories at Time 1 periods. The significant propensity scoring predictors that demonstrated program and comparison youth came from the same overall sample of gang youth were self-reported gang membership, previous experiences with authority, prior total arrests, self-reported drug-using, and perception of non-gang crime in their neighborhoods. The same cross-site predictors were also significant in matching youth within sites. However, statistical significance levels were higher (better) across sites. We are currently comparing different types of coordination and services by program workers to determine which best predict reduction of different arrests and combinations of arrests. Systematic, detailed data, regularly obtained from outreach youth workers/treatment/case managers/school personnel, along with data from police, probation, and parole workers were gathered for this purpose.

Spergel, Irving A., Kwai Ming Wa, and Rolando V. Sosa. 2006. The comprehensive community-wide gang program model: Success and failure. In (James F. Short, Jr. and Lorine A. Hughes, eds.), *Studying youth gangs*. Lanham, MD: Altamira.

Sullivan, Mercer L. 2006. Are "gang" studies dangerous? Youth violence, local context, and the problem of reification. In (James Short, Jr. and Lorine A. Hughes, eds.), *Studying youth gangs*. Lanham, MD: Altamira.

U.S. Government Accountability Office (GAO). 2009. *Combating gangs: Better coordination and performance measurement would help clarify roles of federal agencies and strengthen assessment of efforts*. Washington, DC: Author.

Van Gemert, Frank, Dana Peterson, and Inger-Lise Lien. 2008. *Street gangs, migration and ethnicity*. Devon, UK: Willan.

Statutes Cited

Gang Abatement and Prevention Act of 2007, S. 456, 110th Cong. (1997).

Gang Abatement and Prevention Act of 2009, S. 132, 111th Cong. (2009).

Gang Prevention, Intervention and Suppression Act of 2009, H.R. 1022, 111th Cong. (2009).

Youth Prison Reduction through Opportunities, Mentoring, Intervention, Support, and Education Act (or the Youth Promise Act of 2007), H.R. 3846, 110th Cong. (2007).

Youth Prison Reduction through Opportunities, Mentoring, Intervention, Support, and Education Act of 2009, H.R. 1064, 111th Cong. (2009).

Irving Spergel's interest in gang research developed while working with gang youths from 1952 to 1960, as part of the New York City Youth Board. The board was one of the earliest organized responses to youth gangs. Through its street-worker program, the organization worked to reduce gang tensions by building long-term relationships with members and by providing constructive, alternative activities. Throughout the span of his career, Spergel has written more than 100 books, articles, monographs, and other publications. Perhaps the best known among them are *Racketville, Slumtown, Haulburg: An exploratory study of delinquent subcultures*, and 1995's, *The youth gang problem: A community approach*—a leading work in the field. A major contribution of Spergel's work—known as the "Spergel Model"—is its call for a coordinated approach that encompasses law enforcement, community groups, schools, social service agencies, and governmental organizations. The U.S. Department of Justice has endorsed the Spergel Model, testing the program in more than 20 cities around the country during the last decade. The National Youth Gang Center uses the model as a central component in its work. A test of the model in Little Village Chicago revealed a 40% reduction in serious violence for the 200 program youths, compared with an equivalent sample of nonparticipating youths from the same gangs during a 5-year period.

RESEARCH ARTICLE

GANG DATABASES

Gangs and public policy
Constructing and deconstructing gang databases

Julie Barrows
University of Minnesota

C. Ronald Huff
University of California, Irvine

Research Summary

Attention to gang issues has dramatically increased in the last several decades, both in the scholarly literature and in law enforcement. Despite widespread attention to the gang problem, researchers, police officers, and lawmakers have yet to agree on definitions used to characterize and understand the problem. This article summarizes the existing literature concerning the importance of accurately defining and classifying gang members, documents and analyzes state and federal gang legislation in the United States, and provides a detailed analysis of one state's system that might serve as a useful model for other states.

Policy Implications

Serious risks to public safety and civil liberties are associated with Type 1 and Type 2 classification errors regarding gang membership. The wide variation in state statutory definitions of "gang member" and in the construction and administration of gang databases presents major challenges for policymakers and academic researchers. This article addresses these challenges and argues that a more rigorous and unified system, based on one state's existing model, might be possible and could offer significant advantages in our efforts to address the delinquent and criminal behavior of gangs throughout the United States.

Keywords

gangs, gang member identification, gang databases, gang intelligence systems

During the last several decades, gangs and gang-related crime and delinquency have attracted increasing attention from researchers and policymakers. According to the latest National Youth Gang Survey, U.S. law-enforcement agencies estimated 788,000 active gang members and 27,000 gangs, and they reported that the trends in gang-related homicides and other violent offenses had increased significantly during the previous year (Egley and O'Donnell, 2009).[1],[2] Based on the Tenth Amendment to the U.S. Constitution, also known as the States' Rights Amendment, the prevention, intervention, investigation, prosecution, and punishment of gangs and gang members rests squarely on the shoulders of state governments.[3] How do states respond to this responsibility? How do different jurisdictions classify individuals as gang members? Are these classifications trustworthy? What are the implications of such classification systems? Toch (2007) compared gang classification systems in prisons with medieval persecutions of witches, arguing that depending on "reliable" sources for gang membership identification, "guilt by association," and self-admission, among other criteria, is reminiscent of primitive inquisitions. Can gang classification systems be antithetical to due process? Are such systems necessary to promote public safety?

Bjerregaard (2003) examined anti-gang legislation, particularly the California Street Terrorism Enforcement and Prevention (STEP) Act, and argued for improved definitions and legislation governing gangs and gang members. Drawing on the work of McCorkle and Miethe (2002), who determined that the (then) newly enacted Nevada gang legislation was scarcely employed, Bjerregaard concluded that it is possible for gang legislation that is not carefully crafted to "backfire and actually exacerbate the [gang] problem" (Bjerregaard, 2003: 184).

In his case study of a Midwestern gang unit (described more fully below), Katz (2003) also found that gang legislation, particularly gang enhancement legislation, was severely underused. Additionally, he found that the official policies of the gang unit were largely ignored. For example, a gang in Junction City is not documented pursuant to official definitions. Rather, gang officers rely on a criteria-based system for classifying gang members first, before documenting the gang to which the members belong. Furthermore, gang officers "do not verify that the gang meets the official definition of the gang unit" (Katz, 2003: 502).

---

1.   The magnitude of gang-motivated crime and delinquency is significantly lower than gang-related offenses. The former requires evidence of gang motivation in committing the offense, whereas the latter requires only a correlation—an offense was committed and a gang member was involved.

2.   Larger cities and suburban counties accounted for 60% of the reported gangs and 80% of the reported gang members, although the estimated number of gang members in the larger cities declined by 0.7%. The largest increases in estimates of gangs and gang members occurred in smaller cities and rural counties.

3.   See Appendix A for a listing of state and federal legislation cited herein.

The contributions by Bjerregaard (2003), McCorkle and Miethe (2002), and Katz (2003) signal several issues pertinent to this article in the area of gang definitions, gang policies, and gang legislation. First, law-enforcement agencies may have informal and formal policies for defining and maintaining information on gangs and gang members. Second, these policies may or may not be consistent with legislation. Third, and most important, there can be a great divide between that which is written—whether in local policy or state legislation—and that which is actually practiced by gang enforcement officers. In an effort to narrow this divide, we examine what is written in state codes.

This article provides a comprehensive analysis and critique of state legislation concerning gang definitions and gang databases, as well as the criteria used to classify gang members. We begin with a discussion of why clear definitions and accurate databases are necessary. We examine how states define gangs and classify gang members. After this broad overview, we offer a detailed analysis of one state's approach that could be a useful model for other states. Finally, we discuss some recent federal efforts to address the gang problem. Ultimately, we argue that under the ideals of cooperative federalism, improvements in gang definitions and treatment of gang information is possible and perhaps necessary.

Denial, Overreaction, and Misidentification

Huff (1989, 1990) characterized the typical community's treatment of a gang problem as a three-stage process including denial, overreaction, and misidentification. The initial stage, wherein the presence of a gang problem is denied, has often plagued gang prevention and enforcement efforts. After this initial stage, the typical reaction (the overreaction stage) is to "define the problem as 'a law enforcement problem'" (Huff, 1990: 312). This is analogous to treating cancer as solely a "physicians' problem." Typically occurring after a highly publicized violent gang crime, policymakers implement an enforcement strategy. The problem is that police often overreact, conducting indiscriminate "sweeps," for example, and classifying marginal gang associates, who are often situated at an elevated level of risk because they either (1) do not understand or abide by the "rules" of the gang, (2) attempt to "prove" themselves to the gang, or (3) do not understand the consequences of gang membership—any of which alone or in concert could put them in harm's way. More central to this discussion, many if not most law-enforcement agencies include marginal gang associates in their databases. Indeed, it will be shown that police often include in a database individuals who meet only one criterion of gang membership (other than self-admission, which is generally a reliable indicator, as noted by Curry, 2000 and Webb, Katz, and Decker, 2006), without the requirement of a criminal conviction or even an arrest. Huff (1992: 27) explained, "one of the dangers to be avoided by police is overreacting by prematurely classifying ('labeling') these marginal 'wannabes' as gang members and targeting them for arrest."

The third and related, stage is "misidentification" (Huff, 1990), wherein police misidentify (i.e., overidentify) gang members and the causes of gang membership. The result of misidenti-

fication is twofold. First, police often target the wrong individuals, thereby potentially driving them into gang membership because they are being treated as, and known as, gang members, thus exacerbating the crime problem, including violent crime. Second, researchers and policymakers have inadequate information with which to understand gang membership or guide policy. The policies of police clearly need to be reviewed to ensure an appropriate balance in the identification of gang members—a balance that should neither underidentify nor overidentify individuals as gang members.

The Need for Accurate Identification

The implications of incorrectly classifying individuals as (non)gang members are consequential and can be dangerous. The consequences of a Type 1 error, whereby a non-gang member is categorized as a gang member, are alarming. First, law enforcement may give undue attention to such an individual. Second, if a non-gang-involved person is placed on a list of members of a particular gang, and if rival gang members discover that the individual is believed to associate with that particular gang, then the rival gang members could conceivably target that individual for violence (perhaps causing that person to join a gang for perceived "protection"). Third, it is also conceivable that the non-gang member could be suspended or expelled from school, terminated from employment (see Katz, 2003), or judged unfairly in a civil matter, such as a tenant–landlord dispute. Indeed, some states have legislation targeting gang members in these areas (see Washington, Arkansas, and Arizona, respectively). Finally, consistent with labeling theory (Bernburg, Krohn, and Rivera, 2006; Lemert, 1967), this individual might be so ostracized from normative institutions that the identity of "gang member" is embraced (secondary deviance).

A bona fide gang member who is not identified as a gang member (i.e., a Type 2 error) poses a serious risk to society. A great deal of research has shown that gang members are far more criminally active and violent than are non-gang members (Battin-Pearson, Thornberry, Hawkins, and Krohn, 1998; Decker and Van Winkle, 1996; Howell and Decker, 1999; Huff, 1996, 1998, 2004; Klein, 1995; Thornberry and Burch, 1997; Thornberry, Krohn, Lizotte, Smith, and Tobin, 2003). Failure to identify a gang member could result in the gang member's enhanced ability to commit crime because of less awareness by law enforcement. Furthermore, many states have legislation that enhances penalties for committing a crime to benefit a gang (see National Youth Gang Center, 2006a). The analysis of such gang enhancement legislation is reserved for necessary future research. For our purpose here, it is important to note that if an undocumented gang member is apprehended and charged with a crime, that individual might not receive the appropriate sanctions. Finally, the undocumented gang member might not be referred to an appropriate intervention program.

Thus, gang members clearly need to be identified accurately. The challenge to law enforcement is that gangs are typically characterized by fluidity in membership and differential organizational structures (Weisel, 2002). The geographic mobility of gang members (Maxson,

1998) poses additional challenges to accurate identification, especially in light of the differing classification systems employed by various law-enforcement jurisdictions. Law enforcement has sought to overcome this latter issue by creating databases to house gang information.

Constructing the Gang Database

Gang databases have appeared in law-enforcement jurisdictions across the United States. The need for maintaining and sharing gang information is perhaps unparalleled in this nation's history, largely because of the increased mobility of gang members combined with technological advances, which allow for electronic storage and dissemination. Yet, researchers predicted this need many years ago. As Huff (1989: 532) stated, "the implications for law enforcement are clear: to effectively contain these gang-related offenses, police must have some centralized unit or, at the very least, must share intelligence on gang members and their activities." More recently, Wilson (1997) provided an overview of intelligence information unique to gangs, such as the gang moniker (i.e., the street name assigned to gang members by their peers), and urged law-enforcement agencies to collect and share such information in a database. He recommended that criteria for classifying an individual as a gang member be established in written policy and warned investigators to check whether their state statutes regulate this type of information. He also recommended that an agency coordinator only be allowed to modify information contained in a gang database, but he also warned that the coordinator should not be the same individual who is responsible for auditing the system. Overall, Wilson's guidelines are straightforward and many agencies have followed his advice. However, neighboring jurisdictions compile gang information according to their own gang definitions and criteria, which results in the potential for inconsistency in information from one gang database to the next. If databases are not unified and gang definitions continue to vary, there is the potential for defense attorneys to question the validity of information contained in one database because such information would not be allowed or included in another gang database (Wilson, 1997). To prevent such a scenario, he concluded that "because gangs do not respect law enforcement jurisdictional boundaries, a unified, multi-jurisdictional database would provide the optimum response to gang-related crime" (1997: 17).

Gang Databases in the Literature

Gang databases, or gang intelligence systems, have undergone some scrutiny by gang researchers. As Katz, Webb, and Schaefer (2000) noted, some scholars—such as Chesney-Lind, Rockhill, Marker, and Reyes (1994); Hagedorn (1990); and Spergel (1995)—are pessimistic about the validity and usefulness of gang information. The concern centers on police discretion and the potential failure to accurately distinguish gang-involved from non-gang-involved individuals. In an effort to test the validity of gang intelligence systems, Katz et al. (2000) used data from the Mesa (Arizona) Police Department's Gang Unit and the Maricopa County (Arizona) Juvenile Probation Department, and compared the criminal histories of documented gang members and associates with non-documented delinquent youth. Contrary to the findings of Zatz (1985,

1987) and McCorkle and Miethe (1998), they found that documented gang members and associates were far more criminally active than their undocumented delinquent counterparts. Specifically, documented gang members were "significantly more likely to have been arrested for violent felony offenses (42.5% compared to 18.9%)" than were documented gang associates, who were in turn "significantly more likely to have been arrested for violent offenses (45.3% compared to 28.5%)" than were non-gang delinquent youth (Katz, et al., 2000: 429). The authors concluded that "these findings suggest that gang information systems may be more helpful to the police than first believed" (Katz, et al., 2000: 432).

In addition to aiding gang investigations, another benefit of the gang database is that collecting such information could have a positive impact on a gang-involved youth's life. Huff (1998: 7) described the first "window of opportunity" for gang intervention as "the year between the 'wannabe' stage and the age at first arrest." Although police are not likely to have contact with an individual at this stage, contact during the second window of opportunity is guaranteed. Huff (2004) described the second window as the period between a gang member's first offense, which is usually a minor property offense, and subsequent offenses, which may become more serious. Thus, if police agencies that are collecting information on suspected gang members were to share such information with intervention agencies, our intervention efforts could be improved, thereby reducing recidivism and harm to victims.

More central to this discussion, Katz et al. (2000: 434) suggest that additional efforts be undertaken to publicize police policies for categorizing gang members. They argued that "open policies regarding documentation practices may improve police–community relations" and that such openness could contribute to the improvement of such policies. They suggested that future research focus on the collection, processing, and dissemination of gang information. More recently, Wright (2006) raised constitutional issues associated with gang databases, arguing that inclusion in such databases increases the likelihood of conviction and the imposition of sentencing enhancements, thus raising Fifth and Fourteenth Amendment issues. Furthermore, Weisel and Howell (2007) questioned the inclusion of juveniles in gang databases and instead recommended that North Carolina not include juveniles in its statewide GangNET system, pending a review of compliance with federal regulations on criminal intelligence data (see 28 C.F.R. § 23 [2008]); if included, they should be separated from the names of adults in the database.

Implications

Before examining current state and federal legislation on gang definitions and databases, it is important to highlight the implications of accurate gang identification. Although the discussion above makes clear the implications of accurate gang identification for law enforcement, prosecutors, and defense attorneys, it is also important to consider the implications for corrections and for communities as a whole.

Gangs in Prison

In their comprehensive examination of gangs in U.S. correctional facilities, Knox et al. (1997) studied more than 3,500 self-reported gang members comprising more than one third of approximately 10,000 convicts in 85 institutions located in 17 states. These researchers found that gang members were more predatory, more violent, more involved in drug dealing, more likely to come from dysfunctional families, more likely to fire a gun at a police officer, and more likely to pose a greater threat to public safety. More recently, Gaes, Wallace, Gilman, Kein-Saffran, and Suppa (2001) and Griffin and Hepburn (2006) determined that gang-involved prison inmates are more likely than non-gang-involved inmates to be violent, even when individual characteristics, such as age, race, criminal history, and security risk, were controlled (in fact, Gaes et al. argued that they are twice as likely). Thus, like members of street gangs, prison gang members pose a greater danger to safety and security than those not affiliated with a gang.

That said, the scope of the problem is unclear. For example, in a survey of 160 juvenile correctional facilities, Knox (1991) found that "fifty-two percent of the responding institutions reported that more than 10 percent of confined youth were involved in gangs" (as cited in Howell, 2004: 22–23). Leiter (1993), however, reported that administrators of juvenile detention halls and juvenile correctional facilities estimated that approximately 40% of those incarcerated were gang involved (as cited in Howell, 2004). Finally, in their survey of 800 incarcerated juveniles in six facilities, Sheley and Wright (1993, 1995) found that "two-thirds (68 percent) of the inmates self-reported affiliation with a gang or a 'quasi-gang'" (as cited in Howell, 2004: 23). So, which is it: 10%, 40%, or 68%? These findings paint dramatically different pictures of the extent of the problem. Are the discrepancies the result of the empirical studies being conducted in different locations at different times with different methodologies? Or did the use of varying gang classification systems have an effect?

In an effort to clarify the scope of the gang problem in prisons, Knox (2005) surveyed security threat group (STG) coordinators in every correctional facility in the United States, resulting in a sample of 193 coordinators in 49 states. The coordinators estimated that 25.9% of males and 6.28% of females entered the institution as gang members and that an additional 11.6% of males and 3.7% of females joined a gang while incarcerated. Thus, it is clear that correctional facilities are fertile grounds for gang recruitment because one in nine male inmates joins a gang while incarcerated. As for criminal activity while in prison, the study found that gang members accounted for 20.6% of the institutional management problems and 26.3% of the violence among inmates. Finally, one of the top priorities of the coordinators was to improve intelligence systems to monitor and track gang members and their activities in prison (Knox, 2005).

Noting the lack of empirical evidence on gangs in jails rather than prisons, Ruddell, Decker, and Egley (2006) analyzed survey responses from jail administrators of 134 jails in 39 states and found that approximately 13.2% of jail inmates were gang involved. Questioning how jail administrators classify inmates as gang involved, Ruddell et al. (2006) provided respondents with five criteria indicative of gang membership and asked whether the jail employed those

particular criteria in their classification system. The five criteria are (1) tattoos, colors, or symbols of membership (used by 86.8% of respondents); (2) membership designated by another law-enforcement agency (83.3%); (3) inmate self-reports membership (81.7%); (4) inmate associates or has been arrested with members (71.4%); and (5) identified by a reliable informant (65.3%). Importantly, Ruddell et al. (2006: 39) noted the "considerable overlap between police and jail methods of classification." Finally, in terms of disruptive behavior, Ruddell et al. (2006) found that gang members were more likely than any other "special-needs" group to assault other inmates.

In sum, gangs do exist, if not flourish, in the confines of correctional facilities. The implication is that corrections officials need to know as much as possible about the inmates. Although incarcerated individuals have different (i.e., fewer) rights than ordinary citizens, and there is a difference between street and prison gangs, there is also a great need for consistency in the application of gang classification systems. Indeed, Ruddell et al. (2006:38) found that "designation of gang membership by another law enforcement agency was commonly used to define gang membership." The street gang label, then, can follow an individual into the correctional facility. A gang or threat group classification system should be established to prevent gang recruitment and violence within correctional facilities.

Community Response

The idea of gang classification systems causes concern among various interest groups. A key objection is that law enforcement is afforded too much discretion in classifying individuals as gang members. Indeed, it will be shown that associating with gang members is often a criterion of gang membership. Does this mean that if an individual has a gang member for a brother or a neighbor, that individual could also be targeted as a gang member? It certainly is possible. Where should law enforcement draw the line?

Katz (2003) conducted an important ethnographic study of the gang data collection, maintenance, and dissemination procedures of Junction City's gang unit. He noted that the gang unit officers, after receiving and reviewing the information about subjects' gang status from patrol officers untrained in gang identification methods, were conservative in their assignment of the gang member designation. Gang officers typically neglected to include on the gang list subjects who, by policy, could have been categorized as "associate/wanna be" members (Katz, 2003). This tendency could result in a dramatic underrepresentation of gang membership, the implication of which was detailed above. Furthermore, and of greater concern here, it underscores the wide discretion police officers have in applying the gang member label. Junction City seemed to have progressive formal policies in place, but they failed to adhere consistently to those policies in practice (Katz, 2003). This begs the question of what happens in the field when Junction City is not under such keen observation. It also raises the question of whether Junction City's gang unit is representative of other gang units.

This highlights the need for accurate and systematic methods for classifying an individual as a gang member. Similar to civil gang injunctions and ordinances (see Maxson, Hennigan,

and Sloane, 2005), gang databases have been at the center of heated debate. For example, the Youth Justice Coalition, which is based in Los Angeles and is a member of the National Juvenile Justice Network, proclaimed on its Web site (2006): "Eliminate the 'War on Gangs!'… Develop practices that end the overuse and abuse of the gang database system." On another Web site, the American Civil Liberties Union (ACLU, 2005), described the racial disparity in the Gang Reporting, Evaluation, and Tracking (G.R.E.A.T.) system (which later became CALGANG), wherein 90% of the 20,221 gang subjects in the database were Latino, Asian, or African American—even though these groups comprised less than half of the overall population. The ACLU asked the county district attorney's office to establish a civilian oversight board to monitor what it saw as problems with the list. Finally, the New Urban Perspective (2004) published an article on its Web site describing how activists feel about the gang database used by the Denver Police Department. Claiming that the database is discriminatory, the article asserts that "In 1993, the database included nearly 7,000 names of gang members between the ages of 12 and 24 and their affiliates. More than half of the list was made up of African-Americans and Latinos." The article then compared these figures with city census data, noting that "the numbers implied that 2 out of every 3 African-American males in Denver was a gang member. For Latinos it was 1 out of 4." The ramifications of incorrectly classifying gang members are many. Classification methods should be based on behavior, not on race or ethnicity. Law enforcement must assure interest groups that gang classification is not a discriminatory practice. This can, in part, be accomplished by demystifying classification policies.

Deconstructing the Gang Database

Scope
Many local agencies create their own databases for storing gang information. In their study of 149 police departments and 191 prosecutors' offices across the nation, Johnson, Webster, Connors, and Saenz (1995) found that 70% of police departments and 20% of prosecutors' offices used an automated system for storing gang information. Additionally, of the police departments that reported a gang problem, 78% used a database (Johnson et al., 1995). This study, which was published nearly 15 years ago, illustrates the growing trend of automating gang intelligence information.

State Gang Statutes and Definitional Criteria
Because it is clear that law-enforcement agencies use gang databases, we examined current state and federal legislation governing gangs. State statutes were reviewed and separated into those providing (1) a gang definition, (2) gang criteria, and (3) guidelines for a gang database. This work draws from and updates the compilation of legislation by the National Youth Gang Center (2006a) and the Gang Intelligence Strategy Committee (2007).

In all, 41 states and the District of Columbia provide statutory definitions of a "gang." Table 1 provides a breakdown of each state by the number of participants required to be considered a gang. Five states do not specify how many individuals must participate (although

Arizona requires at least one member who meets the criteria of gang membership), 33 states require three or more members, three states require at least five members, and only the District of Columbia requires at least six members. All states require that criminal activity is necessary to define a gang; however, there is minor variation on whether a pattern of criminal activity must exist or whether isolated criminal acts define a gang. Furthermore, 26 states require that a gang have a common name or symbol, and only 5 states require an identifiable leadership or hierarchy. One state (Maryland) requires either a hierarchy or a common symbol. Five states require that a gang have members with an understanding or alliance with each other. Overall, only five general elements are involved in defining gangs: number of participants, criminal activity, hierarchy, alliance or understanding, and a common name or symbol.

TABLE 1

Statutes on the Definition of "Gang" (*N* = 41)

State	Criminal Activity	Pattern of Criminal Activity	Specified Criminal Activity	Hierarchy	Alliance/ Conspiracy/ Understanding	Common Name/Symbol
Unspecified number of individuals						
AZ	X		X			
CT	X					
NV	X		X		X	X
OR	X					X
WA	X			X		X
Requires three or more individuals						
AK	X	X	X			X
AL	X	X			X	
AR	X	X	X			
CA	X	X	X			X
CO	X	X	X			
DE	X	X	X			X
FL	X	X	X			X
GA	X	X	X			X
IA	X	X				X
ID	X	X	X			X
IL	X	X	X	X	X	
IN	X		X			
KS	X	X	X			X
LA	X	X	X			X
MA	X					X
MD	X	X		X		X
MN	X	X	X			X
MO	X	X	X			X
MS	X		X	X	X	

MT	x	x	x			x
NC	x	x	x			x
ND	x	x	x			
NH	x	x	x			x
NJ	x		x			
NY	x	x				x
OH	x	x	x			x
RI	x	x				x
SD	x	x	x			x
TN	x	x	x			
TX	x	x		x		x
UT	x	x	x			x
VA	x	x	x			x
WI	x	x	x			x
Requires five or more individuals						
KY	x	x	x	x	x	
SC	x	x				
OK	x		x			
Requires six or more individuals						
DC	x	x	x			

Notes. AZ requires at least one gang member in the group, MD requires a hierarchy or common name or symbol of the gang, and DC requires a pattern of or a specified criminal activity.

Although many states have policies for classifying an individual as a gang member, most do not write the policies into statute. Only 15 states define "gang member" by state statute, and 5 of these states employ a relatively general definition. Illinois and Mississippi define a gang member as an individual

> who actually and in fact belongs to a gang and any person who knowingly acts in the capacity of an agent for or accessory to, or is legally accountable for, or voluntarily associates himself with a course or pattern of gang-related activity, whether in a preparatory, executor, or cover-up phase of any activity, or who knowingly performs, aids, or abets any such activity (740 Ill. Comp. Stat. Ann. 147/10 [2002]; Miss. Code Ann. § 97-44-3 [1999]).

New York and South Carolina similarly define a gang member as an individual who is part or an active member of a gang, respectively (N.Y. Comp. Codes R. & Regs. tit. 9, § 301.3 [2007]; S.C. Code Ann. § 16-8-230 [2007]). Wisconsin defines a gang member as an individual "who participates in criminal gang activity" (Wis. Stat. Ann. § 939.22 [2008]). The definitions employed by these five states, although written into statute, are inadequate. The

ambiguity inherent in, and variable interpretation afforded by, these definitions could result in overidentification. We believe that employing a criteria-based classification method is far more robust. Ten states provide the criteria used to categorize an individual as a gang member in statute. Table 2 displays the states that employ criteria by statute, how many criteria points are necessary to indicate that the individual is a gang member, and a description of each criterion of gang membership.

TABLE 2									

Gang Criteria by State (*N* = 10)

	AZ	FL	ID	KS	NH	NJ	SD	TN	TX	VA
Required Criteria	2	2	2	3	2	2	2	2	2	1
Total Criteria	7	11	6	8	4	7	7	7	6	3
Criteria										
Admits membership	X	X	X	X	X	X	X	X	X	X
Identified by anyone			X							
Identified by parent or guardian		X		X				X		
Identified by reliable source		X		X	X		X	X	X	
Corroborated information from source of untested reliability		X		X			X	X	X	
Testimony or official statement	X					X				
Correspondence	X					X				
Corresponds about crime committed by gang		X								
Resides or frequents gang area and adopts dress, hand signs or tattoos and associates with members			X	X	X		X	X		X
Resides or frequents gang area and associates with members									X	
Frequents gang area and adopts dress, hand signs, tattoos, or symbols									X	
Adopts style of dress	X	X				X				
Adopts hand signs		X								
Tattoos	X	X				X				
Any other indicia	X					X				
Associates with members		X								
Stopped with members 2 or more times				X						
Stopped with members 4 or more times		X	X				X			
Arrested once with members									X	
Arrested more than once with members			X	X	X		X	X		X
Documented with paraphernalia or photograph	X	X	X	X		X	X	X		
Has participated in a gang initiation ritual				X						

Note. KS requires an admission or at least three of the other criteria of gang membership.

A careful examination of Table 2 suggests a few important observations. First, only two states (Arizona and New Jersey) employ the same definition of "gang member." Thus, a gang member in Florida, for example, may not be documented as a gang member in Texas. The ramifications of this have been discussed and are enormous. Second, these ten states agree that self-admission is a criterion of gang membership. Indeed, this is the only criterion used by all ten states. It is the most robust method for determining gang status and is used most commonly in the gang literature as well. However, Kansas uses self-admission as the *sole* criterion necessary for the gang member designation. Potential exists here, too, for overidentification, either through possible intimidation by police to obtain an admission of membership or through an individual's desire to identify with a gang when the gang itself may not confirm that association. Third, the criteria of gang membership employed by the states, taken as a whole, are similar. In examining these criteria, it is apparent that states mimic each other in language and purpose. Given the similarity in criteria, it might be possible for states to develop a consistent gang classification system.

Gang Database Legislation

Finally, we analyzed the state legislation surrounding gang databases. Nine states (Arizona, Colorado, Florida, Illinois, Minnesota, South Carolina, Tennessee, Texas, and Virginia) have statutes governing the compilation, maintenance, and dissemination of gang information. Some other states have legislation hinting at the collection of gang information but, after examining such legislation, it became clear to us that states need to review their policies for creating and maintaining gang databases. Texas seems to have the most robust statute concerning gang databases, and it requires adherence to 28 C.F.R. § 23 (2008), the Criminal Intelligence Systems Operating Policies. Florida repealed its serious habitual juvenile offender database and replaced it with a statewide criminal gang database. However, Florida (and Illinois) fails to provide crucial details, such as procedures for entry, audit, and deletion of records. South Carolina also recently added legislation governing the implementation of a new statewide gang database, which must follow the requirements of the federal gang database discussed below. The other states (Arizona, Colorado, Tennessee, and Virginia) simply do not provide enough information.

The Present Study

The examination of gang legislation demonstrates the need to improve and enhance such legislation. Overall, 41 states provide a gang definition, although all states should do so. Only ten states provide the criteria used to determine gang membership, and just nine states outline guidelines for establishing and maintaining a gang database. Much improvement is necessary and might not be difficult to achieve. Much of the legislation that exists is similar in language and purpose. State legislation should continue to be evaluated in an effort to achieve consensus on gang definitions, criteria of gang membership, and guidelines for gang databases. Based on our comparison of the approaches of various states, we believe that the statutes of Minnesota could serve as a prototype that could help states reach such a consensus. Furthermore, although

Minnesota has not yet committed to statute the criteria of gang membership, the existing criteria are similar to those used in other states and thus could also be used as the starting point in striving for consensus.

The Minnesota Model

It is important to describe the "Minnesota Model" in some detail so that its applicability can be assessed by others.[4] The Metro Gang Strike Force (MGSF) and its Oversight Council adopted a definition of "criminal gang," which was subsequently written into Minnesota's legislative statute (Minn. Stat. § 609.229 [2006]) as follows:

> 609.229 Crime committed for benefit of gang.
>
> Subdivision 1.Definition.
>
> As used in this section, "criminal gang" means any ongoing organization, association, or group of three or more persons, whether formal or informal, that:
>
> (1) has, as one of its primary activities, the commission of one or more of the offenses listed in section 609.11, subdivision 9[5];
>
> (2) has a common name or common identifying sign or symbol; and
>
> (3) includes members who individually or collectively engage in or have engaged in a pattern of criminal activity.

Minnesota does not include a definition of "gang member" in its legislative code. However, the MGSF implemented three definitions that essentially became the state standard. The MGSF differentiates among suspected, confirmed, and confirmed and convicted gang members.

A *suspected* gang member is an individual who has engaged in criminal activity and who has been investigated and observed meeting one or more documented gang criteria according

---

4. We note that there have been some criticisms of Minnesota's Metro Gang Strike Force (MGSF). However, our focus here is on the overall approach represented by the "Minnesota Model" and is not meant to imply that each component of that model functions perfectly in practice.

5. "Minnesota Statute 609.11 Subd. 9. Applicable offenses. The crimes for which mandatory minimum sentences shall be served as provided in this section are: murder in the first, second, or third degree; assault in the first, second, or third degree; burglary; kidnapping; false imprisonment; manslaughter in the first or second degree; aggravated robbery; simple robbery; first-degree or aggravated first-degree witness tampering; criminal sexual conduct under the circumstances described in sections 609.342, subdivision 1, clauses (a) to (f); 609.343, subdivision 1, clauses (a) to (f); and 609.344, subdivision 1, clauses (a) to (e) and (h) to (j); escape from custody; arson in the first, second, or third degree; drive-by shooting under section 609.66, subdivision 1e; harassment and stalking under section 609.749, subdivision 3, clause (3); possession or other unlawful use of a firearm in violation of section 609.165, subdivision 1b, or 624.713, subdivision 1, clause (2), a felony violation of chapter 152; or any attempt to commit any of these offenses."

to the following 10-point gang criteria adopted statewide by the Minnesota Gang Oversight Council:[6]

1. Admits gang membership or association

2. Is observed to associate on a regular basis with known gang members

3. Has tattoos indicating gang membership

4. Wears gang symbols to identify with a specific gang

5. Is in a photograph with known gang members and/or using gang-related hand signs

6. Name is on a gang document, hit list, or gang-related graffiti

7. Is identified as a gang member by a reliable source

8. Arrested in the company of identified gang members or associates

9. Corresponds with known gang members or writes and/or receives correspondence about gang activities

10. Writes about gangs (graffiti) on walls, books, and paper

If an ongoing criminal investigation exists on a subject and the subject meets at least one of the above criteria, then a gang file may be created to include a hard copy of any supporting documentation and an electronic copy to be housed in GangNET, which is a confidential structured query language Web-based database offered by SRI International, Inc. (Menlo Park, CA). Documentation means a report, photograph, or other evidence that can be obtained and reproduced to prove the observation of gang activity.

A *confirmed* gang member is an individual who has engaged in criminal activity but has not necessarily been convicted or adjudicated guilty of a gross misdemeanor or felony. The individual must meet at least three of the ten criteria of gang membership listed above. Prosecutors will use this definition to satisfy the definition of "gang member" for purposes of Minn. Stat. § 609.229 (2006).[7]

A *confirmed and convicted* gang member is an individual who meets the confirmed gang member definition and has been convicted or adjudicated of a gross misdemeanor or felony.

---

6. These criteria are applied to one who has engaged in criminal activity and is believed to have met one or more of these criteria. This approach does not, however, prove gang motivation and might be based on only one of these criteria. This category is somewhat akin to a hypothesis that must be tested with more rigorous criteria.

7. Again, meeting three or more of these criteria does not prove gang motivation. Minnesota uses this only to show gang membership. If the individual continues to commit crime but no longer meets three or more of these criteria, it would no longer be considered gang-related crime, since there would be no link to the gang.

As long as the individual is 14 years of age or older, this subject meets the requirements to be included in a second database called the Minnesota Gang Pointer File, which is part of the state's Criminal Justice Information System (CJIS), maintained by the Minnesota Bureau of Criminal Apprehension. The Pointer File was written into Minnesota legislative statute (Minn. Stat. § 299C.091 [2006]) as:

299C.091 Criminal gang investigative data system.

Subdivision 1.Establishment.

The bureau shall administer and maintain a computerized criminal gang investigative data system for the purpose of assisting criminal justice agencies in the investigation and prosecution of criminal activity by gang members. The system consists of data on individuals whom law enforcement agencies determine are or may be engaged in criminal gang activity. Notwithstanding section 260B.171, subdivision 5, data on adults and juveniles in the system and data documenting an entry in the system may be maintained together. Data in the system must be submitted and maintained as provided in this section.

Subd. 2.Entry of data into system.

(a) A law enforcement agency may submit data on an individual to the criminal gang investigative data system only if the agency obtains and maintains the documentation required under this subdivision. Documentation may include data obtained from other criminal justice agencies, provided that a record of all of the documentation required under paragraph (b) is maintained by the agency that submits the data to the bureau. Data maintained by a law enforcement agency to document an entry in the system are confidential data on individuals as defined in section 13.02, subdivision 3, but may be released to criminal justice agencies.

(b) A law enforcement agency may submit data on an individual to the bureau for inclusion in the system if the individual is 14 years of age or older and the agency has documented that:

(1) the individual has met at least three of the criteria or identifying characteristics of gang membership developed by the Gang and Drug Oversight Council under section 299A.641, subdivision 3, clause (7), as required by the council; and

(2) the individual has been convicted of a gross misdemeanor or felony or has been adjudicated or has a stayed adjudication as a juvenile for an offense that would be a gross misdemeanor or felony if committed by an adult.

Subd. 3.Classification of data in system. Data in the criminal gang investigative data system are confidential data on individuals as defined in section 13.02, sub-

division 3, but are accessible to law enforcement agencies and may be released to the criminal justice agencies.

CJIS, which is the overarching computer system hosting the Pointer File, also contains the state's criminal history, missing persons, stolen and recovered vehicles, and other "hot" files that are communicated to the National Crime Information Center (NCIC), which is a division of the Federal Bureau of Investigation (FBI). Information in the Pointer File, however, is not shared with the NCIC. The database is, rather, an officer-safety system accessible to Minnesota peace officers, even from a patrol car. As such, officers who have stopped an individual can query the Pointer File using the subject's name and date of birth to determine whether the subject is a confirmed and convicted gang member. If so, the Pointer File will indicate with which gang the subject is affiliated and whether there are officer-safety issues of concern.

In sum, Minnesota maintains two statewide gang databases: GangNET and the Pointer File. GangNET is a comprehensive database with unlimited capability to store name, date of birth, physical description, moniker, gang affiliation, documented gang criteria, school or employment, residence, vehicle, field interview, as well as arrest and conviction information. Furthermore, photographs and other documents may be uploaded to the system. In addition to the information on suspected, confirmed, and confirmed and convicted gang members, Gang-NET houses information on the gangs identified in Minnesota, including details of the gang's common symbols, beliefs, membership, and "turf" boundaries. As of April 2009, GangNET contained information on 17,408 total gang personnel in 583 gangs.

Rather than attempting to maintain hard files and rely on human memory to track the status of gang subjects, MGSF officials decided to have a systematic and electronic method for tracking the age, criteria, and conviction information of gang subjects. After all, the Pointer File is mandated by the Minnesota Legislature and, as such, a system of controls must be in place to ensure the accuracy of entries into the Pointer File. Likewise, to ensure quality entries into the Pointer File, GangNET must adhere to a system of controls.

Law-enforcement officers who have identified a suspected gang member and have collected associated documentation of gang criteria can enter the information directly into GangNET, provided that (1) their supervisor has reviewed and approved the information and (2) they have received training on entering information into GangNET by the MGSF gang analyst responsible for the system. Alternatively, untrained law-enforcement officers may forward their supervisors' approved hard copy file to the analyst for entry into GangNET. After verifying that the information in the hard file is accurate and represents credible gang criteria documentation, the gang subject may be entered into GangNET. If the individual is already in GangNET, then that individual's electronic and hard file is updated with the new information.

Because the MGSF receives no federal funds to support the operation of GangNET, information could remain in GangNET indefinitely. However, the MGSF adheres to all other regulations set forth under 28 C.F.R. § 23 (2008) in the administration of GangNET. The MGSF gang analyst supervises a team of data entry personnel who consistently audit the information contained in the system. Each instance of a new entry requires the auditing of

the entire subject file. Furthermore, the hard file of a gang member eligible for entry into the Pointer File is audited. Finally, when the Pointer File is audited or purged, GangNET is also audited because the entry in GangNET supports the entry in the Pointer File. It should also be noted that files are purged from GangNET if the entry is determined to be inaccurate or if the gang subject has escaped gang membership.

Like GangNET, the Pointer File uses no federal funds. However, Pointer File administration is also compliant with 28 C.F.R. § 23 (2008). Pointer File audits are legislatively mandated as follows:

> 299C.091 Criminal gang investigative data system.
>
> Subd. 4. Audit of data submitted to system. The bureau shall conduct periodic random audits of data under subdivision 2 that documents inclusion of an individual in the criminal gang investigative data system for the purpose of determining the validity, completeness, and accuracy of data submitted to the system. The bureau has access to the documenting data for purposes of conducting an audit.

The MGSF audits the Pointer File twice a year, and each audit is facilitated by the MGSF analyst. Most Pointer File entries are completed by the staff of the MGSF analyst. However, a few agencies have entry access because an individual within those agencies has been trained on Pointer File entry by the MGSF analyst. Ten percent of the Pointer File entries, or a minimum of two entries, are randomly selected from each entering agency. The MGSF analyst audits the MGSF entries, and the analyst sends an affidavit to all other entering agencies to be signed by a commander or supervisor once the agency's entries have been audited. Inaccuracies are noted (included in an audit report) and are either corrected in or purged from the system.

> Guidelines for purging the Pointer File are legislatively mandated as follows:
>
> 299C.091 Criminal gang investigative data system.
>
> Subd. 5. Removal of data from system. Notwithstanding section 138.17, the bureau shall destroy data entered into the system when 3 years have elapsed since the data were entered into the system, except as otherwise provided in this subdivision. If the bureau has information that the individual has been convicted as an adult, or has been adjudicated or has a stayed adjudication as a juvenile for an offense that would be a crime if committed by an adult, since entry of the data into the system, the data must be maintained until 3 years have elapsed since the last record of a conviction or adjudication or stayed adjudication of the individual. Upon request of the law enforcement agency that submitted data to the system, the bureau shall destroy the data regardless of whether 3 years have elapsed since the data were entered into the system.

Whereas 28 C.F.R. § 23 (2008) allows data to remain in the system for 5 years, the MGSF employs a more stringent maximum of 3 years. A report is printed every month listing the gang members who are eligible to be purged in the following month. The list of gang members is forwarded to the respective entering agency to confirm that the individual has no new conviction or adjudication and thus should be purged from the system. The MGSF analyst then completes the purge. If, however, the gang member has a new conviction or adjudication, then the file is renewed as of the date of the conviction or adjudication and is eligible to remain in the system for an additional 3 years. The data, however, are retained in a hard file as well as in GangNET. If a gang subject is convicted or adjudicated at some point in the future, then the subject may be reentered into the Pointer File.

In our judgment, the Minnesota guidelines for operating GangNET and the Pointer File provide a useful benchmark in the effort to collect, maintain, audit, and purge appropriately sensitive information on gangs and gang subjects. Without a doubt, crimes committed by gang members pose a significant threat to public safety, and law-enforcement agencies need this type of information to protect public safety and solve cases. Indeed, GangNET has helped solve cases. For example, a rape victim described a tattoo on the forearm of her attacker. A description of the tattoo was queried in GangNET, a photo lineup of suspects was created, the victim positively identified her attacker, and he was subsequently convicted (K. Navara, personal communication, October 28, 2008). Likewise, the Pointer File has helped many law-enforcement officers better prepare for their encounters with confirmed and convicted gang members. That is the way these data systems are intended to work, although some areas are in need of improvement.

First, against the advice of Wilson (1997), only one MGSF gang analyst is responsible for overseeing the gang database entry, audit, and purge functions. The legitimacy of the databases will increase with a more objective review of the systems. Second, the technology could be improved to reduce human error and provide easier linking and information sharing to other databases. Third, communication among interested criminal justice professionals should be improved. Many gang investigators are not aware of the gang member definitions used by the MGSF or the MGSF databases that house gang information. Many law-enforcement agencies employ their own definition of a gang member and house information according to their own methods, which often include a separate database. This issue is not unique to Minnesota (see, for example, Katz, 2003). Rather, this issue is highlighted to illustrate that communication and consistency need to be improved nationwide. Researchers and criminal justice professionals should strive for a more unified approach.

A Federal Gang Database: Prospects and Problems

In a statement before the Senate Committee on the Judiciary, the (then) Chief of the Violent Crimes and Major Offenders Section of the FBI, Steven R. Wiley, described the increasing threat of criminal gang activity. After reviewing current studies, he stated that "the current increase in gang activity, including migration into previously gang-free communities, has required

federal, state, and local law enforcement agencies to adjust resources to deal with the resulting increase in violent crimes and drug trafficking" (Wiley, 1997). He then described how federal law enforcement has developed new strategies to combat the growing problem, including the creation of the Violent Gang and Terrorist Organization File (VGTOF) on October 1, 1995, and stated that "the database provides identifying information about gangs and gang members to law enforcement personnel. This information serves to warn law enforcement personnel of the potential danger posed by violent individuals, and promotes the exchange of information about gangs and their members to facilitate criminal investigations" (Wiley, 1997).

In our judgment, although the idea of a federal gang database is appealing, there are three major problems in the implementation of the VGTOF. First, the federal government does not impose its definition of a criminal street gang on state or local government. Although new legislation is pending under the proposed Gang Abatement and Prevention Act of 2007, the federal government currently defines a criminal street gang (18 U.S.C. § 521 [2009]) as:

Title 18 Part I Chapter 26 § 521 Criminal street gangs

(a) Definitions.

"conviction" includes a finding, under state or federal law, that a person has committed an act of juvenile delinquency involving a violent or controlled substances felony.

"criminal street gang" means an ongoing group, club, organization, or association of 5 or more persons-

that has as 1 of its primary purposes the commission of 1 or more of the criminal offenses described in subsection (c);

the members of which engage, or have engaged within the past 5 years, in a continuing series of offenses described in subsection (c); and

the activities of which affect interstate or foreign commerce.

"State" means a state of the United States, the District of Columbia, and any commonwealth, territory, or possession of the United States.

This definition, which requires that a gang have five or more members, is more stringent than most state codes, which typically require only three members. Thus, because the federal government requires a membership of at least five individuals, and because the federal government does not impose this definition on state and local government, the VGTOF is detached from the methods employed by state and local government and therefore is not helpful.

Second, gang definitions vary from state to state and even from city to city. Neither researchers nor law-enforcement personnel have reached consensus on a common definition of "gang" or "gang member." Gang researchers have long noted this issue in the literature (see Curry, Ball, and

Fox, 1994; Esbensen, Winfree, He, and Taylor, 2004; Horowitz, 1990; Howell, 2000; Howell and Gleason, 1999; Klein, 1995; Klein, Gordon, and Maxson, 1986; Knox, 2002; Maxson, 1998). Moreover, various jurisdictions operationalize the term differently. In other words, although agencies may have a formal definition of "gang," informal—often discretionary—methods are employed for classifying an individual as a gang member. For these reasons, many law-enforcement agencies have created their own databases for storing information on gangs and gang members and are often reluctant or unwilling to share their data with agencies using different operational definitions. Given the varying gang definitions and the creation of local gang databases, providing local information on gang members to the VGTOF may be viewed as irrelevant on a national level as well as duplicative in terms of data entry effort.

Finally, the VGTOF contains information on gangs and terrorists in the same file. Because of their significant qualitative differences, the value in storing information on gang members and terrorists (and other security threat groups, such as prison gangs, bikers, or adult criminal gangs) in a combined file is not apparent to us and reinforces harmful misconceptions. Furthermore, the focus of the VGTOF has shifted to terrorism, largely because of the events of September 11, 2001. As a result, more questions are now surfacing about the relevance of asking local agencies to store their local gang information in such a file.

Conclusions and Recommendations

Like Thomas ("Tip") O'Neill's famous axiom that "all politics is local," gang problems are often unique to specific locales. An individual in one jurisdiction may be considered a gang member and a viable threat to public safety, whereas the same individual may not be considered a gang member or a threat in another jurisdiction. Therefore, local agencies need to collect and store information pertinent to their respective jurisdictions. However, gang members also cross jurisdictional boundaries. To protect public and officer safety, information should be shared across jurisdictions. More work should be done to bridge various databases. What is the best way to store gang information? Is one unified national database, or even an electronic bridge to connect countless local and regional databases, the answer? What are the implications if various jurisdictions and researchers have differing methods for classifying individuals as gang members?

Although some researchers downplay the importance of agreeing on one definition (see, for example, Horowitz, 1990), others view it as a necessity. Howell (2004: 42) concluded that "a key issue in combating youth gangs is providing a uniform definition for them." Yet, what should that definition be? According to Esbensen et al. (2004), gang identification through self-nomination promises to be the best measure of gang membership for researchers and for police. For researchers, although the size of the gang problem varies using the self-report criterion, Esbensen et al. (2004) argued that the demographic characteristics of gang members remain stable. In other words, researchers would be hard pressed to determine an accurate *extent* of the gang problem, but they could make enlightening conclusions about its *nature*. These authors (2004: 70) explained, "the validity of the self-nomination method lends credence to the police practice of targeting youth who claim gang affiliation," and they encouraged the police

to remove former members from their gang lists because failure to do so results in overidenti-fication, which artificially inflates the magnitude of the problem. It is also important to note that Klein, Kemer, Maxon, and Weitekamp (2001), working with the Eurogang consortium, demonstrated that consensus on an operational gang definition might be possible, even across nations with different cultures and languages (Decker and Weerman, 2005). However, more research is needed to determine whether one operational definition "holds up" across many more nations. Van Gemert, Peterson, and Lien (2008) have addressed this issue and have highlighted the importance of issues such as culture, racism, segregation, and poverty with respect to the use of a definition.

Miller (1990) claimed that one reason why the nation has not solved the gang problem is that the federal government has neglected antigang efforts in its domain. He argued that "without a center of responsibility at the federal level— an office with the authority and resources to develop a national strategy and monitor efforts at the national, state, and local levels—the enterprise of gang control is severely handicapped" (Miller, 1990: 277). More than a decade after Miller's comments, Wesley D. McBride, who is the influential president of the California Gang Investigators Association and a 35-year veteran of the Los Angeles County Sheriff's Department, claimed in his statement before the Senate Committee on the Judiciary (2003: 26) that "there has been no federal leadership in the world of gang enforcement. Gang enforcement still tends to be done by pockets of investigators with little or no communication between these isolated pockets." Thus, researchers and criminal justice practitioners alike argue for an increased federal initiative.

The National Gang Intelligence Center (NGIC) represents one response to the call for increased federal assistance:

> The NGIC will enable the FBI and its local, state, and federal partners to central-ize and coordinate the national collection of intelligence on gangs in the U.S., and then analyze, share, and disseminate this intelligence with law enforcement authorities throughout the country. The NGIC will give local, state, and federal investigators and intelligence analysts the opportunity and mechanism to share their collective information and intelligence on gangs. This will enable gang in-vestigators and analysts to identify links between gangs and gang investigations, to further identify gangs and gang members, to learn the full scope of their criminal activities and enterprises, to determine which gangs pose the greatest threat to the U.S., to identify trends in gang activity and migration, and to guide them in coordinating their investigations and prosecutions to disrupt and dismantle gangs. The NGIC will be an essential part of our efforts to combat and dismantle gangs and will enhance the existing liaison and coordination efforts of federal, state, and local agencies (Swecker, 2005).

Likewise, the National Gang Targeting, Enforcement & Coordination Center (Gang TECC)

was implemented by the U.S. Attorney General in 2006 as a national gang task force working closely with the NGIC. Jaffe (2008) described the Gang TECC and the NGIC as "one-stop-shops" for gang investigative assistance and intelligence information. The extent to which these new related national agencies will help solve the gang problem remains to be observed. Much skepticism is expected, given the different responsibilities and experiences of federal, state, and local law-enforcement agencies and the distrust that has often characterized their interactions. However, it could represent an important first step toward a serious consideration of a nationally coordinated gang intelligence database and a potentially valuable federal effort to assist concerned local agencies and the communities they serve.

Finally, we believe that the nation's interests would be well served at this time if a commission (somewhat like the President's Commission on Law Enforcement and the Administration of Justice in the 1960s) could be formed to analyze current data and research on gangs, assess the interjurisdictional issues presented by gang-related crime, and determine whether one operational definition might be agreed upon to be used in one shared, nationwide database. The significant mobility of gang members is now recognized as a major challenge for cities and states relying solely on local or (only partially) national databases. The degree of readiness to undertake such discussions may be higher than at any time in recent memory. This effort could be coordinated by the National Institute of Justice, led by the Attorney General with his emphasis on its importance, and its membership should ensure that the interests of all relevant parties are represented in the discussions. Such a commission would represent a viable response to the criticisms of Miller (1990) and others that the federal government has abdicated its responsibility with respect to the prevention and control of gang-related crime.

References

American Civil Liberties Union. 2005. *Racial profiling*. Retrieved August 19, 2006 from aclu.org/racialjustice/racialprofiling/19901res20050815.html.

Battin-Pearson, Sara R., Terence P. Thornberry, J. David Hawkins, and Marvin D. Krohn. 1998. *Gang membership, delinquent peers, and delinquent behavior*. Juvenile Justice Bulletin. Rockville, MD: Office of Juvenile Justice and Delinquency Program.

Bernburg, Jon Gunnar, Marvin D. Krohn, and Craig J. Rivera. 2006. Official labeling, criminal embeddedness, and subsequent delinquency: A longitudinal test of labeling theory. *Journal of Research in Crime and Delinquency*, 43: 67–88.

Bjerregaard, Beth. 2003. Antigang legislation and its potential impact: The promises and pitfalls. *Criminal Justice Policy Review*, 14: 171–192.

Chesney-Lind, Meda, Anna Rockhill, Nancy Marker, and Howard Reyes. 1994. Gangs and delinquency: Exploring police estimates of gang membership. *Crime, Law, and Social Change*, 21: 201–228.

Curry, G. David. 2000. Self-reported gang involvement and officially recorded delinquency. *Criminology*, 38: 1253–1274.

Curry, G. David, Richard A. Ball, and Robert J. Fox. 1994. *Gang crime and law enforcement record keeping* (Research in brief). Washington, DC: National Institute of Justice.

Decker, Scott H. and Frank M. Weerman (eds.). 2005. *European street gangs and troublesome youth groups*. Lanham, MD: Alta Mira.

Decker, Scott H. and Barrik Van Winkle. 1996. *Life in the gang: Family, friends, and violence*. New York: Cambridge University Press.

Egley, Arlen, Jr. and Christina E. O'Donnell. 2009. *Highlights of the 2007 National Youth Gang Survey*. Washington, DC: U.S. Department of Justice, Office of Justice Programs.

Esbensen, Finn-Aage, L. Thomas Winfree, Jr., Ni He, and Terrance J. Taylor. 2004. Youth gangs and definitional issues: When is a gang a gang, and why does it matter? In (Finn-Aage Esbensen, Stephen G. Tibbetts, and Larry Gaines, eds.), *American youth gangs at the millennium*. Long Grove, IL: Waveland Press.

Gaes, Gerald G., Susan Wallace, Evan Gilman, Jody Klein-Saffran, and Sharon Suppa. 2001. *The influence of prison gang affiliation on violence and other prison misconduct*. Washington, DC: Federal Bureau of Prisons.

Gang Intelligence Strategy Committee. 2007. *Brief review of federal and state definitions of the terms "gang," "gang crime," and "gang member."* Washington, DC: U.S. Department of Justice, Office of Justice Programs, Global Justice Information Sharing Initiative, Global Intelligence Working Group.

Griffin, Marie L. and John R. Hepburn. 2006. The effect of gang affiliation on violent misconduct among inmates during the early years of confinement. *Criminal Justice and Behavior*, 33: 419–466.

Hagedorn, John M. 1990. Back in the field again: Gang research in the nineties. In (C. Ronald Huff, ed.), *Gangs in America*. Newbury Park, CA: Sage.

Horowitz, Ruth. 1990. Sociological perspectives on gangs: Conflicting definitions and concepts. In (C. Ronald Huff, ed.), *Gangs in America*. Newbury Park, CA: Sage.

Howell, James C. 2000. *Youth gang programs and strategies*. Washington, DC: Office of Juvenile Justice and Delinquency Prevention.

Howell, James C. 2004. Youth gangs: An overview. In (Finn-Aage Esbensen, Stephen G. Tibbetts, and Larry Gaines, eds.), *American youth gangs at the millennium*. Long Grove, IL: Waveland Press.

Howell, James C. and Scott H. Decker. 1999. *The youth gangs, drugs, and violence connection* (Juvenile Justice Bulletin). Washington, DC: Office of Juvenile Justice and Delinquency Program.

Howell, James C. and Debra K. Gleason. 1999. *Youth gang drug trafficking. juvenile justice bulletin*. Washington, DC: Office of Juvenile Justice and Delinquency Prevention.

Huff, C. Ronald. 1989. Youth gangs and public policy. *Crime & Delinquency*, 35: 524–537.

Huff, C. Ronald. 1990. Denial, overreaction, and misidentification: A postscript on public policy. In (C. Ronald Huff, ed.), *Gangs in America*. Newbury Park, CA: Sage.

Huff, C. Ronald. 1992. The new youth gangs: Social policy and malignant neglect. In (Ira M. Schwartz, ed.), *Juvenile justice and public policy: Toward a national agenda*. New York: Lexington Books.

Huff, C. Ronald. 1996. The criminal behavior of gang members and nongang at-risk youth. In (C. Ronald Huff, ed.), *Gangs in America*, 2nd Edition. Thousand Oaks, CA: Sage.

Huff, C. Ronald. 1998. *Comparing the criminal behavior of gang members and at-risk youths* (Research in brief). Washington, DC: National Institute of Justice.

Huff, C. Ronald. 2004. Comparing the criminal behavior of gang members and at-risk youths. In (Finn-Aage Esbensen, Stephen G. Tibbetts, and Larry Gaines, eds.), *American youth gangs at the millennium.* Long Grove, IL: Waveland Press, Inc.

Jaffe, David. 2008. The Department of Justice's gang squad, gang targeting enforcement and coordination center, and national gang intelligence center. *United States Attorneys' Bulletin*, 56: 8–9.

Johnson, Claire M., Barbara A. Webster, Edward F. Connors, and Diana J. Saenz. 1995. Gang enforcement problems and strategies: National survey findings. *Journal of Gang Research*, 3: 1–18.

Katz, Charles M. 2003. Issues in the production and dissemination of gang statistics: An ethnographic study of a large Midwestern police gang unit. *Crime & Delinquency*, 49: 485–516.

Katz, Charles M., Vincent J. Webb, and David R. Schaefer. 2000. The validity of police gang intelligence lists: Examining differences in delinquency between documented gang members and nondocumented delinquent youth. *Police Quarterly*, 3: 413–437.

Klein, Malcolm W. 1995. *The American street gang*. New York: Oxford University Press.

Klein, Malcolm W., Margaret A. Gordon, and Cheryl L. Maxson. 1986. The impact of police investigations on police-reported rates of gang and nongang homicides. *Criminology*, 24: 489–512.

Klein, Malcolm W., Hans-Juergen Kerner, Cheryl L. Maxson, and Elmar G. M. Weitekamp. 2001. *The Eurogang paradox: Street gangs and youth groups in the U.S. and Europe*. Dordrecht, The Netherlands: Kluwer.

Knox, George W. 1991. *An introduction to gangs*. Barren Springs, MI: Vande Vere Publishing.

Knox, George W. 2002. Responding to gangs in the 21st century: A research and policy review. *Journal of Gang Research*, 9: 63–74.

Knox, George W. 2005. *The problem of gangs and security threat groups (STGs) in American prisons today: Recent research findings from the 2004 prison gang survey*. Retrieved September 4, 2006 from ngcrc.com/corr2006.html.

Knox, George W., Jodet-Marie Harris, Thomas F. McCurrie, Alice P. Franklin Elder, Edward D. Tromanhauser, John A. Laskey, et al. 1997. A gang classification system for corrections: Some results of the 1996 Project GANGFACT study: A large comparative analysis of 3,500 gang members in 17 states. *Journal of Gang Research*, 4: 47–57.

Leiter, Valerie. 1993. *Special analysis of data from the OJJDP Conditions of Confinement Study*. Cambridge, MA: Abt Associates.

Lemert, Edwin M. 1967. *Human deviance, social problems, and social control*. Englewood Cliffs, NJ: Prentice-Hall.

Maxson, Cheryl L. 1998. *Gang members on the move*. Juvenile Justice Bulletin. Washington, DC: Office of Juvenile Justice and Delinquency Prevention.

Maxson, Cheryl L., Karen M. Hennigan, and David C. Sloane. 2005. It's getting crazy out there: Can a civil gang injunction change a community? *Criminology & Public Policy*, 4: 577–605.

McBride, Wesley D. 2003. *Combating gang violence in America: Examining effective federal, state, and local law enforcement strategies*. Hearing before the Committee on the Judiciary, United States Senate.

McCorkle, Richard C. and Terance D. Miethe. 1998. The political and organizational response to gangs: An examination of a 'moral panic' in Nevada. *Justice Quarterly*, 15: 41–64.

McCorkle, Richard C. and Terance D. Miethe. 2002. *Panic: The social construction of the street gang problem*. Upper Saddle River, NJ: Prentice Hall.

Miller, Walter B. 1990. Why the United States has failed to solve its youth gang problem. In (C. Ronald Huff, ed.), *Gangs in America*. Newbury Park, CA: Sage.

National Youth Gang Center. 2006a. *Compilation of gang-related legislation*. Retrieved October 30. 2008 from iir.com/NYGC/gang-legis/.

New Urban Perspective. 2004. *Activists attack gang database*. Retrieved August 19, 2006 from newurbanperspective.org.

Ruddell, Rick, Scott H. Decker, and Arlen Egley, Jr. 2006. Gang interventions in jails: A national analysis. *Criminal Justice Review,* 31: 33–46.

Sheley, Joseph F. and James D. Wright. 1993. *Gun acquisition and possession in selected juvenile samples* (Research in brief). Washington, DC: Office of Juvenile Justice and Delinquency Prevention.

Sheley, Joseph F. and James D. Wright. 1995. *In the line of fire: Youth, guns, and violence in urban America.* Hawthorne, NY: Aldine de Gruyter.

Spergel, Irving. 1995. *The youth gang problem.* New York: Oxford University Press.

Swecker, Chris. 2005. *FBI efforts to combat gangs with ties to Central America and Mexico*. Statement before the Subcommittee on the Western Hemisphere, House International Relations Committee. Retrieved October 30, 2008 from ciponline.org/facts/050420swec.pdf.

Thornberry, Terence P. and James H. Burch, II. 1997. *Gang members and delinquent behavior*. Juvenile Justice Bulletin. Washington, DC: Office of Juvenile Justice and Delinquency Program.

Thornberry, Terence P., Marvin D. Krohn, Alan J. Lizotte, Carolyn A. Smith, and Kimberly Tobin. 2003. *Gangs and delinquency in developmental perspective*. Cambridge, UK: Cambridge University Press.

Toch, Hans. 2007. Sequestering gang members, burning witches, and subverting due process. *Criminal Justice and Behavior,* 34: 274–288.

Van Gemert, Frank, Dana Peterson, and Inger-Lise Lien. 2008. *Street gangs, migration, and ethnicity.* Devon, UK: Willan.

Webb, Vincent J., Charles M. Katz, and Scott H. Decker. 2006. Assessing the validity of self-reports by gang members: Results from the Arrestee Drug Abuse Monitoring Program. *Crime & Delinquency,* 52: 232–252.

Weisel, Deborah Lamm. 2002. *Contemporary gangs: An organizational analysis.* New York: LFB Scholarly.

Weisel, Deborah Lamm and James C. Howell. 2007. *Comprehensive gang assessment: A report to the Durham Police Department and Durham County Sheriff's Office.* Durham, NC: Durham Police Department.

Wiley, Steven R. 1997. *Violent street gangs in America.* Statement before the Senate Committee on the Judiciary. Retrieved October 30, 2008 from hi-ho.ne.jp/taku77/ refer/gang.htm.

Wilson, Craig R. 1997. What's in a name? Gang monikers. *FBI Law Enforcement Bulletin,* 66: 14–17.

Wright, James D. 2006. The constitutional failure of gang databases. *Stanford Journal of Civil Rights & Civil Liberties*, 2: 115–142.

Youth Justice Coalition. 2006. "Eliminate the 'War on Gangs.'" Retrieved August 19, 2006 from njjn.org/members_public_state_CA.html.

Zatz, Marjorie. 1985. Los Cholos: Legal processing of Chicano gang members. *Social Problems*, 33: 13–30.

Zatz, Marjorie. 1987. Chicano youth gangs and crime: The creation of a moral panic. *Contemporary Crises*, 11: 129–158.

Julie Barrows is a Ph.D. candidate in sociology at the University of Minnesota. Her dissertation focuses on the formation, network structure, and effectiveness of multi-agency gang task forces. In addition to her academic training, she has worked in the criminal justice system for more than 10 years, including experience at every level of government and at every major juncture along the "criminal justice funnel." Her interest in responses to gangs stems from nearly 5 years of experience as a crime analyst for the Metro Gang Strike Force in Minnesota.

C. Ronald Huff, Ph.D., is dean of the School of Social Ecology at the University of California, Irvine and professor of criminology, law, and society and of sociology. He previously taught at Ohio State University, where he directed both the John Glenn School of Public Affairs and the Criminal Justice Research Center; Purdue University; and the University of Hawaii. His research focuses on gangs, wrongful conviction, and public policy. His publications include more than 80 journal articles and book chapters as well as 12 books, the most recent of which are *Wrongful conviction: International perspectives on miscarriages of justice* (Temple University Press, 2008) and Gangs in America III (Sage, 2002). He is a fellow and past president of the American Society of Criminology. His other honors include an Outstanding Academic Book Award from the American Library Association (for *Convicted but innocent: Wrongful conviction and public policy*); the Donald Cressey Award from The National Council on Crime and Delinquency; the Paul Tappan Award from the Western Society of Criminology; and the Herbert Bloch Award from the American Society of Criminology. He has served as a consultant to the attorneys general of California, Ohio, and Hawaii; the U.S. Senate Judiciary Committee; the FBI National Academy; the U.S. Department of Justice; and numerous other federal, state, and local agencies throughout the nation.

A P P E N D I X   A

Table of Authorities

State Statutes

Ala. Code § 13A-6-26 (2005) Compelling Streetgang Membership
Alaska Stat. § 11.81.900 (2008) Definitions
Ariz. Rev. Stat. Ann § 13-105 (2001) Definitions
Ariz. Rev. Stat. Ann §13-3829 (2001) Criminal Street Gang and Sex Offender Monitoring; Qualified Vendor; Definition
Ariz. Rev. Stat. Ann. § 33-1368 (2001) Noncompliance with Rental Agreements by Tenant; Failure to Pay Rent; Utility Discontinua-
    tion; Liability for Guests; Definition
Ark. Code Ann. § 5-74-103 (2005) Definitions
Ark. Code Ann. § 21-15-111 (2004) Hiring New Employees into Designated Financial or Information Technology Positions
Cal. Penal Code § 186.22 (2008) Participation in Criminal Street Gang; Penalty
Colo. Rev. Stat. § 18-23-101 (2004) Definitions
Colo. Rev. Stat. § 24-33.5-415.3 (2008) Information on Gangs — Legislative Declaration
Conn. Gen. Stat. § 29-7n (2003) Record and Classification of Gang-Related Crimes
D.C. Code § 22-951 (2005) Criminal Street Gangs
Del. Code Ann. Tit. 11, § 616 (2007) Gang Participation
Fla. Stat. § 874.03 (2000) Definitions
Fla. Stat. § 874.09 (2000) Crime Data Information
Ga. Code Ann. § 16-15-3 (2007) Definitions
Idaho Code Ann. § 18-8502 (2004) Criminal Gang Enforcement Act
20 Ill. Comp. Stat. Ann. 2640/10 (2008) Statewide Organized Gang Database Act
740 Ill. Comp. Stat. Ann. 147/10 (2002) Illinois Streetgang Terrorism Omnibus Prevention Act
Ind. Code § 35-45-9 (2004) Criminal Gang Control
Iowa Code § 723A.1 (2003) Definitions
Kan. Stat. Ann. § 21-4226 (2006) Definitions
Ky. Rev. Stat. Ann. § 506.140 (1999) Criminal Gang Recruitment — Definitions for Chapter
La. Rev. Stat. Ann. § 15:1403 (2005) Criminal Street Gangs and Patterns of Criminal Street Gang Activity; Prohibitions and Criminal
    Penalties
Mass. Gen. Laws ch. 265, § 44 (2002) Coercion of Child Under Eighteen into Criminal Conspiracy; Penalties
Md. Code Ann., Criminal Law § 9-801 (2008) Definitions
Minn. Stat. § 299C.091 (2006) Criminal Gang Investigative Data System
Minn. Stat. § 609.229 (2006) Crime Committed for Benefit of a Gang
Miss. Code Ann. § 97-44-3 (1999) Mississippi Streetgang Act Definitions
Mo. Rev. Stat. § 578.421 (2003) Definitions
Mont. Code Ann. § 45-8-402 (2007) Definitions
N.C. Gen. Stat. § 15A-1340.16 (2007) Aggravated and Mitigated Sentences
N.D. Cent. Code § 12.1-06.2 (1997) Criminal Street Gangs
N.H. Rev. Stat. Ann. § 651:6 (2008) Extended Term of Imprisonment
N.J. Stat. Ann. § 2C:33-28 (2008) Crime of Gang Criminality; "Criminal Street Gang" Defined; Grading of Offense
Nev. Rev. Stat. § 213.1263 (2007) Board may Prohibit Association with Members of Criminal Gang as Condition of Parole
N.Y. Comp. Codes R. & Regs. tit. 9, § 301.3 (2007) Types of Activities Not Allowed on State Property
Ohio Rev. Code Ann. § 2923.41 (2008) Criminal Gang Definitions
Okla. Stat. tit. 21, § 856 (2002) Definitions
Or. Rev. Stat. § 336.109 (2003) Policy to Reduce Gang Involvement, Violent Activities and Drug Abuse
R.I. Gen. Laws § 12-19-39 (2002) Criminal Street Gang Enhancement
S.C. Code Ann. § 16-8-230 (2007) Definitions
S.C. Code Ann. § 16-8-330 (2007) Development and Management of Statewide Criminal Gang Database
S.D. Codified Laws § 22-10A-1 (2006) Definitions

Tenn. Code Ann. § 38-6-102 (2008) Criminal Investigation Division – Criminal Intelligence Unit – Investigations into Domestic Terrorism – Emergency Situation Traffic Stops

Tenn. Code Ann. § 40-35-121 (2006) Criminal Gang Offenses – Enhanced Punishment – Procedure

Tex. Code Crim. Proc. Ann. art. 61.01-11 (2006) Criminal Street Gang Intelligence Database

Tex. Code Crim. Proc. Ann. art. 71.01 (2006) Definitions

Utah Code Ann. § 76-9-802 (2008) Definitions

Va. Code Ann. § 18.2-46.1 (2004) Definitions

Va. Code Ann. § 52-8.6 (2005) Criminal Street Gang Reporting

Wash. Admin. Code § 148-120-100 (2006) Conduct Violations

Wash. Rev. Code Ann. § 28A.600.455 (2009) Gang Activity- Suspension or Expulsion

Wis. Stat. § 939.22 (2008) Words and Phrases Defined

Federal Statutes

18 U.S.C. § 521 (2009) Criminal Street Gangs; Definitions

28 C.F.R. § 23 (2008) Criminal Intelligence Systems Operating Policies

U.S. Const. Amend. X

GANG DATABASES

Gang databases

Context and questions

James B. Jacobs
NYU School of Law

Long before computers were so widespread, individual police officers and the gang-intelligence units of big city police departments maintained intelligence files on gangs and gang members. The introduction of computers enabled more systematic information gathering and storage, vastly more powerful search capacity, and practically instantaneous retrieval. Today, databases on gangs and gang members have proliferated at the local, state, and federal levels. As Barrows and Huff (2009, this issue) explain, the trend is clearly in the direction of more databases and access to information on more gang members—whether it is through a centralized gang database or diverse decentralized databases.

Gang databases are just one of many species of criminal justice investigative and intelligence databases that the information technology revolution has made possible. The police and other investigative, prosecutorial, and correctional agencies record many types of information in an electronic form, which can be easily accessed by countless authorized users and then lawfully shared or unlawfully leaked. Once information is entered into an investigative or intelligence database, it can easily migrate to other public and private databases and, therefore, can become more difficult to purge or edit effectively. Mistakes might never be corrected. The label could last a lifetime.

Gang databases come in various forms. Increasingly, police create gang databases for intelligence purposes—independent of conviction, arrest, or even a criminal investigation. In addition, jail and prison officials record such information in their databases in furtherance of institutional order and discipline. Prosecutors and judges also might want access to such information to use in making charging, plea bargaining, and sentencing decisions. In short, criminal justice agencies and agents operate many different databases wholly or partly devoted to gangs and gang members. They also routinely share gang-intelligence information with one other.

Recording gang affiliation in an intelligence database is controversial because intelligence databases are constructed by police agencies on the basis of what might be limited, ambiguous, or even mistaken information. But even gang databases created by prosecutors and by jail and prisons officials are vulnerable to error, because a decision about whether a particular individual is a gang member (and whether a particular group of individuals constitutes a gang) often requires discretionary judgment. Moreover, even if a person is accurately identified as a "gang member," the meaning and significance of such membership varies greatly from individual to individual.

Civil Liberties Concerns

Gang databases raise civil liberties concerns because the concept of "gang" is amorphous. The label can be affixed to a tightly knit group of professional criminals or to friends who hang out together, and to all sorts of "groups" and clusters of individuals in between. Some critics perceive gang labeling as stigmatizing minority inner-city youth and, worse, limiting their educational, social, and economic opportunities as a consequence. The critics emphasize the vast numbers of youth—especially Black and Hispanic youth—who are entered into gang databases.

Despite these concerns, political and legal challenges to gang-intelligence databases have not been nearly as successful as challenges to intelligence initiatives directed at political radicals and suspected terrorists. Recall the 1960s and 1970s controversies about the Federal Bureau of Investigation and police information gathering on "radicals" and political activists. Legal challenges to some of these operations resulted in tight controls as to when and what kind of intelligence files could be opened and what use could be made of the information. Recently, the federal government's intelligence operations in the war on terror have triggered a firestorm of protests. Apparently, that kind of intelligence gathering is much more contentious than intelligence gathering directed at suspected gang members.

Intelligence gathering directed at suspected gang members has more or less been treated as a legitimate law-enforcement strategy rather than as a form of political repression. For decades, the police and other law-enforcement agencies have maintained organized crime files on the La Cosa Nostra (LCN) crime families, their members, hierarchies, activities, residences, businesses, and social clubs. To my knowledge, these files and databases have never been challenged successfully. Undoubtedly, the reason is that LCN is regarded as an unambiguous crime problem. To the extent that gangs are understood to be proto-organized crime groups, intelligence gathering on them is regarded as unproblematic.

Access to Gang Databases

Who has a good reason for accessing a gang intelligence database? The police argue, plausibly, that good information about gangs and gang members is essential for protecting neighborhoods and for properly deploying covert and undercover personnel and youth workers. Jail and prison officials argue that gang information is imperative for making appropriate classification and other decisions directly related to institutional safety and security. Prosecutors argue that knowledge

of a defendant's gang affiliation and rank better informs charging and plea bargaining decisions. Judges argue that a defendant's gang affiliation is relevant to bail and sentencing decisions. School officials argue that knowing which students belong to a gang helps them maintain security and assign counseling resources. Employers argue that screening gang members from the job applicant pool contributes to a reliable and competent workforce that will best promote the interests of the business and its customers and employees. If all these agencies, entities, and individuals have access to gang-intelligence databases, then the information is effectively public.

What About Due Process?

Being recorded in a database as a gang member can have fateful consequences for an individual's interactions with the criminal justice system, schools, employers, and even landlords. Yet, the police, jails, and prisons can label an individual as a gang member without affording him or her any opportunity to contest the label and without notifying the individual that the labeling has occurred. Although lack of notification seems blatantly unfair to the individual, it is hard to see how due process could come to the rescue. It would often defeat the idea of intelligence gathering to notify the target that he or she has been identified as a criminal threat. Imagine notifying suspected members and associates of LCN that they have a right to contest the organized crime-control unit's determination that they might be connected to organized crime!

Even if the intelligence target were to be notified, then what kind of due process would be appropriate? Would the suspected gang member have the right to a hearing before a police commander or some "neutral" fact finder? If so, would the police intelligence gatherers have to reveal to the target what they knew and how they knew it? Would they have to expose their informants, undercover agents, and electronic surveillance? Would the sources of intelligence information be subject to cross-examination? Would the suspected gang member be entitled to representation by retained counsel? Would the state have to provide an attorney to an indigent gang member? Would there be a right to an appeal within the police department or to a court? These questions strongly suggest that establishing a due process right to contest inclusion in a gang database is unlikely. Thus, maintaining the accuracy of gang databases will necessarily depend on the values, judgment, and competence of those officials who control the database.

Centralized or Decentralized Gang Databases?

Is the legislature the appropriate decision maker for authorizing or prohibiting the creation of a gang database? Should one size fit all counties in a state? Gang problems (and gang-control resources) vary considerably from community to community. Generally speaking, law enforcement in the United States is highly decentralized. Each county has its own prosecutor, and each city or town has its own police department. Although legislative attempts to control prosecutorial and police discretion do occur (e.g., mandatory arrest and prosecution in domestic violence cases)—for the most part—policing and prosecutorial policy is set at the local level. Furthermore, it is difficult to enforce state-level policy against local-level criminal justice system officials.

Even if policymaking on gang databases is centralized at the state level, the state is not likely to have a single comprehensive gang database or to tightly control the labeling of gang members. Gang enforcement is a local government function. All gang and gang-member information will originate at the local level and will almost certainly be recorded in local gang-intelligence files and databases. Then, that information will be input directly into the state-level gang database or be transmitted to state-level personnel who will also decide if and how to store the information. Meanwhile, it will probably be possible for local police departments to obtain gang-intelligence information directly from other local police departments.

Neither a state- nor a federal-level gang database can be exclusive. Even if gang databases are combined or merged at a central (e.g., federal or state) level, then it is likely that local police departments would keep their own databases and files (that is certainly true of rap sheets). The age of mainframes has passed; we live in the age of laptops and desktops. A local gang-intelligence officer could maintain his or her own database.

Regulating the Gang Database

Some jurisdictions have sought to regulate gang databases through legislation or administrative rule making. Such laws and rules contain criteria for labeling an individual as a gang member, and include protocols for auditing the database and for purging dated and inaccurate information. Here, willingness, capability, and competence are the relevant issues. Police departments (and other government agencies) always are strapped for resources. Realistically, scrutiny of the gang database is not going to be a high police-department priority. I think it is likely that auditing will be conducted shoddily or not at all. In any event, whether to leave a name in the database will often require subjective judgment. Moreover will information removed from the database be stored in yet another database?

If a name is removed from the gang database, what assurance is there that it will not be reentered in a local police-department database, in a gang-intelligence unit file, or into the computer of an individual police officer? The longer a name remains in a database, the greater the chance that it will be accessed by someone who could enter it into another database, perhaps a private database. The information technology revolution has made it difficult to destroy information effectively. And in the case of gang-intelligence databases, it is unlikely that any "complainant" will hold the feet of the database operator to the fire. Thus, it should not be surprising to find a large discrepancy between how the gang database should operate and how it does operate.

References

Barrows, Julie and C. Ronald Huff. 2009. Gangs and public policy: Constructing and deconstructing gang databases. *Criminology & Public Policy*. This issue.

James B. Jacobs, legal scholar and sociologist, is Warren E. Burger Professor of Law and Director, Center for Research in Crime and Justice, New York University (NYU) School of Law. Among his books are *Mobsters, Unions, & Feds: The Mafia and the American Labor Movement* (NYU Press, 2006) *Gotham Unbound: How New York City was Liberated from the Grip of Organized Crime* (1999), *Busting the Mob: United States v. Cosa Nostra* (1994), and *Corruption and Racketeering in the New York City Construction Industry* (1990), all published by NYU Press.

Gangs and public policy

Constructing and deconstructing gang databases

David M. Kennedy
John Jay College of Criminal Justice

J
ulie Barrows and C. Ronald Huff (2009, this issue) get it right when it comes to the problems with gang database systems. Such databases can be overinclusive, underinclusive, poorly conceived in particular places, inconsistent from place to place, guided by statutes and policies that are themselves all of the foregoing, and often implemented in practice in ways that are inconsistent with these statutes and policies. These discrepancies matter. Our national estimates of gangs and gang membership, through the National Youth Gang Survey, are constructed by adding up local submissions; to the extent that local characterizations of gangs are inconsistent, partial, overreaching, and noncompliant—which they certainly are— our national estimates will be distorted, which they also certainly are. Our understanding of gang crime is based on local designations, leading to such ridiculousness as the Federal Bureau of Investigation's characterization, through the local submission of Supplemental Homicide Reports, that fewer than 1,000 gang homicides occurred in 2005 (Bureau of Justice Statistics, 2007) out of a national total of 16,740; city studies consistently suggest that half or more of all homicides should be attributed to "gangs" (the quotes are used advisedly). Wrongly designat- ing, or not designating, a group as a gang or an individual as a gang member can mean that it and he gets undeserved law-enforcement attention, fails to get deserved attention, does or does not trigger statutory sentencing enhancements, does or does not lead to segregation in jail and prison, as well as causing a host of other consequences. It is not implausible to then suggest that we should have well-founded and consistent definitions, procedures, and systems across jurisdictions.

The so-far insoluble scholarly difficulty of constructing an adequate definition of "gang" should be a clue as to how the task of constructing a legally compliant, conceptually sound, useful, and interjurisdictionally acceptable model for gang databases is likely to go; that is, not well. The scholarly project has foundered on the conceptual problems involved, such as how

to frame a definition that, in its core descriptive elements, does not incorporate those same elements that the definition then will be used to explore causally (e.g., violent offending). Additional considerations are:

- The vast range of phenomena that look, feel, and taste like "gangs" and must then be formally reconciled (e.g., prison gangs, biker gangs, juvenile gangs, and drug crews)
- The differing geographic presentation of "gangs" and their similarities and differences (e.g., Bloods in Compton and Bloods in Newark have much in common and yet are entirely different)
- The wide range of gang behavior (e.g., some gangs have leadership and some do not; some have induction rituals and some do not; and some are territorial and some are not)
- The tendency of definitions to lead to the inclusion of groups that are somewhat similar but, which for hard-to-formalize reasons, look, feel, and taste different (e.g., the Mafia and a feckless local drug crew) and even groups that are really not the same at all (e.g., the Mafia and a fraternity)
- The real presentation of gang offending, in which the causal, instrumental, and hierarchical connections among the gang, the gang offender, and the gang offense can be present, absent, clear, and obscure.

Scholars have been at this issue for generations and have not yet reached a satisfactory resolution. Thus, it seems to be asking a lot for law-enforcement practitioners to succeed where the academy has failed.

The Minnesota system Barrows and Huff (2009) suggest as a model unfortunately is a case in point. The argument made is that the system uses consistent and appropriate definitions of "gang" and "gang member" and has a set of operating, oversight, and audit principles that allow it to both be useful and protect against inaccuracy and over-inclusion. This system makes it a good candidate for fulfilling Barrows and Huff's larger goal that the federal government, state, and local jurisdictions should use a common system. So how might this system work in practice?

I have for the last several years been involved in a gang violence prevention project in Cincinnati called the Cincinnati Initiative to Reduce Violence (CIRV), with colleagues from the University of Cincinnati (UC) and a range of practitioner partners, including the Cincinnati Police Department (CPD). A key element in CIRV has been the systematic identification of what we will call, for the moment, gangs in Cincinnati. Dr. Robin Engel and a team of graduate students from UC's Policing Institute, working with access to CPD data and officers, have conducted repeated systematic audits of the city's gangs. They have identified approximately 60 such gangs; charted their activities, rivalries, and alliances; identified approximately 1,500 individuals in them; run criminal history checks on about half of those individuals—which lead to the finding that the individuals averaged about 35 prior charges apiece, with a third having 10 or more felony charges—and discovered that they are associated as offenders, victims, or both with nearly three quarters of all Cincinnati homicides (i.e., gang members are known to have

been the perpetrators of more than 60% of the city's killings; it is this type of finding, common to such area studies, which makes it clear that the FBI's national numbers on gang killings are ludicrous). The CIRV team also has identified a much smaller number of "impact players" and shooters from the general gang population. This picture is precisely why tracking, understanding, and dealing with gangs is important: They are extraordinarily active and dangerous and typically *are* the homicide and drug problem in many jurisdictions. Cincinnati's gangs, which drive much more than half of all killings, represent 0.3% of the city's population.

The Minnesota system would have missed nearly all of these data. Relatively few of Cincinnati's gangs meet Minnesota's criterion that a gang have "a common name or common identifying sign or symbol"; some do, such as the "Cotti Boys" and the "Taliband" (the recent target of a comprehensive CPD investigation and known locally as "the gang that couldn't spell straight"), but most do not. The CIRV team, which follows the practice of CPD street officers, usually has labeled them simply by street intersections or locations: Green and Race, Green and Republic, and Fay Apartments. These groups are not Bloods, Crips, Latin Kings, Vice Lords, or MS-13; they do not have names and symbols; they usually have little or nothing in the way of leadership; they do not sign; and they do not have induction or exit rituals.

They might not even be gangs. Whatever we decide that term to mean—academically, legally, or in common usage—these groups might not rise to it. In fact, CIRV does not use the term, a decision that was made deliberately because of the freight of meaning that "gang" carries with it and the corresponding distraction that freight causes. CIRV simply refers to groups, which people certainly are, and tracks such things as "GMIs" (group member involved homicides).

This fact pattern is common on the ground. Such offending groups are at the heart of serious violent crime, drug activity, and a wide range of other core public-safety problems. If they are gangs, then the Minnesota system and most other databases need to be expanded to include these groups and, in so doing, will lose the restrictive precision that is one of their key features. If they are not gangs, then a gang database will miss a great deal of what matters. In many places, the core offending groups simply are not, in the minds of both scholars and law enforcement, what is meant when they say "gang." Where they are, that "ganginess" is often largely irrelevant. Similar work in Newark showed groups that claimed Crip and Blood, but formal analysis found little gang coherence along Crip and Blood lines (McGloin, 2005), and qualitative work with knowledgeable police officers showed that the groups had been around for a long time, that they had recently started claiming, that this claiming had virtually no impact on their behavior, and that "gang" conflict was as likely to be between groups that claimed together as groups in nominal conflict. The police officers paid a great deal of attention to these groups and clearly understood their actions and dynamics; for these officers, the groups' Crip and Blood attributes had almost no salience. Mercer Sullivan (2006), in his study of similar issues in New York City at a similar time, found similar dynamics: Many types of offending groups existed; the groups were active; most groups were not what we would call "gangs;" and many that did present in some fashion as gangs, which included claiming Crip and Blood, really were not.

Similar problems develop with protocols that govern what counts as gang membership. The Minnesota system is typical in requiring that one or several particular standards be met. It does not take much work to construct plausible situations in which the younger brother of a gang member could be proved by law enforcement to have met these standards—to have been observed associating on a regular basis with gang members, is in photographs with gang members flashing signs, and has been arrested with gang members—without being a gang, or even a group, member. From the other direction, one of the most dangerous figures in Cincinnati, well known to the police, while a leader in what is an active and dangerous group, is older, wiser, more careful, and probably meets none of the Minnesota criteria.

These problems are intrinsic to definitions of gang and gang membership, are well known and venerable, and are probably insurmountable. They often are used to belittle the idea of gangs: If we cannot figure out a definition that clearly distinguishes a gang and a football team, or a gang member and his mother (who certainly has associated with him and who has probably written him in prison, which would make her a gang member in Minnesota), then obviously gangs do not exist. This concept is nonsense, of course. Football teams and gangs are different, we know they are different, and the fact that we cannot capture that in precise language changes nothing. But it certainly does make taking important steps, which include legal action, based on imprecise language something of a problem.

This issue raises the interesting question of what such systems actually do for us. Barrows and Huff (2009) suggest several plausible reasons why we should care about gang databases. Type 1 errors can lead to profiling, victimization, and collateral damage from schools, employers, and such. Type 2 errors can lead law enforcement to overlook gangs and gang members, raise officer-safety issues, lose the power of legislative gang enhancements, and such. These problems are genuine concerns. It is not clear, however, that we need to solve the gang definition problem and create coherent, far less nationally consistent, gang databases to solve them.

Type 1 errors can be dealt with entirely by not having gang databases. No classification means no possibility of misclassification: Presto, problem solved. Addressing Type 2 errors requires information and vigilance but does not require a gang database. Dangerous groups and dangerous offenders can be tracked perfectly well by ordinary law-enforcement intelligence methods and systems; we do not need to worry about "gangs" to monitor them properly. One of the striking things about projects like Cincinnati's CIRV is the extent to which they demonstrate how well law enforcement does understand what is going on. The initial identification and network analysis of Cincinnati's violent groups was performed by Engel's UC team, which used structured qualitative methods in cooperation with front-line CPD officers, in less than a day. Subsequent work identified individual group members, 71% of whom turned out *already* to be flagged in CPD data systems as officer-safety risks.

Systematically asking and answering such basic questions—which groups and which offenders are the most dangerous and committing the most crime—will be more accurate and more useful if the "gang" question is ignored. The reason projects like CIRV have to go

through such mapping and identification exercises is because gang databases do chronically miss dangerous groups that do not feel like or meet legal criteria to be gangs, and because meticulous criteria and procedures for allowing data in, keeping it in, and scrubbing it invariably, in practice, mean that such systems are out of date, clunky, and incomplete. Few problems occur when we simply bypass this protocol altogether. If gangs and gang members need to be legally certified (e.g., for the application of statutory gang sentencing enhancements), then it can be done when they have drawn investigative attention and have faced prosecution. That is precisely what Cincinnati did when it dismantled the Taliband recently, an investigation that was launched because the group was profligately active and violent, that proceeded operationally on the basis that it was an offending network that needed to be disrupted, that integrated formal social-network analysis by Engel's team (Coolidge, 2009a), and that was dramatically effective (Coolidge, 2009b); and recently won the International Association of Chiefs of Police Excellence in Criminal Investigations Award.

Barrows and Huff (2009) make a good case for enhancing the consistency of gang databases and protocols and for using the Minnesota system as an example, or even a prototype, of so doing. The case holds, though, only if we think we need to track gangs *qua* gangs. In fact, we do not. What Sullivan (2006) called the "reification" of gangs is real, and an enormous distraction. We had no trouble addressing another class of virulent criminal networks, the Mafia, without painstaking and legal definitions of what it meant to be the Mafia and clumsy, statute-governed databases of Mafia members. The Mafia was real, and local and federal authorities knew it. They kept intelligence and mounted investigations, and it worked perfectly well. Even legislation that created special legal regimes aimed at organized crime—like the 1968 Omnibus Crime Control and Safe Streets Act and the 1970 Racketeer-Influenced and Corrupt Organizations Act (RICO)—defined and criminalized certain *conduct*; it did not define a proscribed group or criminalize membership in that group, nor did it criminalize or enhance penalties for acts carried out by that group (the lack of definition of such a group later allowed the RICO statute to be deployed, controversially and ultimately unsuccessfully, against anti-abortion groups). Sullivan titled his article about the mostly-not-gangs he found in New York City as "Maybe we shouldn't study gangs." Maybe, and we probably should not keep databases on them either.

References

Barrows, Julie and C. Ronald Huff. 2009. Gangs and public policy: Constructing and deconstructing gang databases. *Criminology & Public Policy*. This issue.

Bureau of Justice Statistics. 2007. Homicide trends in the U.S. Washington, DC: Department of Justice. Retrieved June 20, 2009 from ojp.usdoj.gov/bjs/homicide/tables/circumsttab.htm.

Coolidge, Sharon. 2009a. Cincinnati police fight crime with computers. *Cincinnati Enquirer*, July 17.

Coolidge, Sharon. 2009b. Social networking helped bust gang. *Cincinnati Enquirer*, July 17.

McGloin, Jean Marie. 2005. Policy intervention considerations of a network analysis of street gangs. *Criminology & Public Policy*, 4: 607–635.

Sullivan, Mercer. 2006. Maybe we shouldn't study gangs: Does reification obscure youth violence? *Journal of Contemporary Criminal Justice*, 21: 170–190.

David M. Kennedy is the director of the Center for Crime Prevention and Control and professor of anthropology at John Jay College of Criminal Justice in New York City. From 1993 through 2004, he was a senior researcher and adjunct professor at the Program in Criminal Justice Policy and Management, Kennedy School of Government, Harvard University. His work focuses on strategies for assisting troubled communities. He has written and consulted extensively in the areas of community- and problem-solving policing, police corruption, and neighborhood revitalization. He has performed field work in police departments and troubled communities in many U.S. and foreign cities. He is the author of a seminal work on community policing, *Deterrence and Crime Prevention: Reconsidering the Prospect of Sanction* (London: Routledge, 2008), and of numerous articles on police management, illicit drug markets, illicit firearm markets, youth violence, and deterrence theory, which include editorials in the *New York Times*, *Washington Post*, and elsewhere.

POLICY ESSAY

GANG DATABASES

Street gang databases

A view from the gang capitol of the United States

Malcolm W. Klein

University of Southern California

Whereas the article on gang databases (Barrows and Huff, 2009, this issue) deals primarily with state legislation and state-level databases, I respond with comments on local police-run gang databases. These suffer seriously from problems of reliability and validity, along with major variations in content. Their value for research purposes is very limited. State-level databases must be derived from aggregated local databases. They cannot overcome the frailties associated with local police data-gathering procedures, and they will suffer as well from the "noise" associated with different local systems. Policymakers turning to such databases to construct state legislation will necessarily be misled as to the nature of street gangs and their control.

Julie Barrows and C. Ronald Huff (2009) have done a commendable job of thoroughly reviewing and dissecting the problems and promises of state-level approaches. At the same time, they have left open the same issues with respect to local, usually police-controlled, databases, so I will concentrate my reaction on this area. Most database use for gang prevention, intervention, law enforcement, and criminal trials is local use. It is helpful to remember that the original authors' 50 or so states contain between 3,000 and 4,000 local jurisdictions reporting on gang problems to the National Youth Gang Center every year. That is a lot of local action.

Telling Examples

I will start by offering just a few examples of gang "knowledge" based on local gang databases and their inclusion criteria. Are girls included in street gang activity? Most police departments in gang-involved cities record 0% to 5% or so female membership. Most ethnographic, observational, and survey research studies report female percentages that range from 10% to 40%.

Why the discrepancy? Several reasons have been suggested:

1.   Police are simply not interested in nor attentive to female gang participation.

2.   Female criminal involvement, although mirroring male patterns of crime, is far lower in the level of crime committed.

3.   As a corollary to the above, lower levels of female violence and Part I crimes make female members less visible to the police who are most concerned with serious forms of gang crime.

4.   Girls desist from gang involvement at an earlier age, thus reducing their vulnerability to police contact and recording.

The result is at least twofold. First, policymakers, often relying on police accounts of gang activity, tend to ignore a substantial fraction of gang membership. Second, criminologists often ignore, avoid, or are disinterested in female gang activity.

My second example comes from the huge gang database that has been maintained for decades by the Los Angeles County Sheriff's Department (LASD), which is charged with collecting and recording gang data from every police department in the county. For many years, the gang unit's report consistently offered a figure of approximately 150,000 gang members throughout the county, based on individual departments' summaries. Then, after the retirement of a key figure in the unit, the sheriff announced at a public seminar that the "real figure" was in the neighborhood of 90,000 gang members. One year later, the figure was reduced to 80,000. Within 2 years, either the county's gang membership was reduced by more than 45% (how is that for effective gang suppression?) or the database was ridiculous. I leave it to the reader to decide.

My third example has to do with the criteria used by law-enforcement agencies to enter an individual into their gang databases. With some variations, most police departments require that a "documented" gang member meet two or three of the following criteria:

•   Self-admission
•   Identification by a "reliable informant"
•   Identification by other informants
•   Use of "gang-style" clothing
•   Display of gang signs or symbols
•   Gang tattoos
•   Frequents gang territories
•   Contacts (by officers) in the company of gang members
•   Arrests with known gang members

Because gang membership is usually a judgmental affair—there is seldom an organizational roster, record of dues-paying membership, or other documented formal gang membership system— criteria such as those above make some sense. But obviously some criteria are more valid than others, and the requirement of meeting two or three of them does little to add certainty. I note, finally, that an additional criterion that used to complete the list was "or any other criteria."

My fourth example comes from the LAPD in 2008. Its database yielded 400 street gangs, 41,000 gang members, and 6,877 gang-related crimes. Leaving aside the questionable definitional problems and exactitude associated with the first two figures, consider just the third statistic: 6,877 gang-related crimes. This figure yields one crime for every six gang members. Readers familiar with self-reported offense data from oral and paper interviews of gang members can only scoff at this minimal level of alleged activity. Why the discrepancy—beyond the obvious fact that most gang crimes are not observed by the police (LAPD or others)? The answers lie in what the LAPD considers worth reporting and recording. Only 11 serious felonies or "gang-style" offenses are of interest.[1] Furthermore, state legislation (Penal Code 186:22) provides dramatically repressive sentences for convicted gang members when it can be shown that their crime was committed "for the benefit of the gang" or "in furtherance of gang goals." Such evidence is not readily available in most instances, and so the police often underreport what cannot lead to enhanced sentences. The database does not reflect crime but crimes convenient to police and prosecutors.

Misidentification Problems

Barrows and Huff (2009) correctly point to both Type 1 and Type 2 errors in state databases. One such problem is the underreporting of female members and crime. Another problem is the omission of short-tenure gang members or inconsistent members whose presence on the street often goes unnoticed and unrecorded by the police. Similarly, younger gang members are far less likely to be visible to the police. Much of this invisibility is "balanced" by the tendencies of the police to (1) over-identify local youth as gang members using some of the questionable criteria listed earlier and (2) maintain members in the database well beyond their behavioral manifestation of membership. Most gang members became ex-gang members, although not in the eyes of many officers and database recorders ("once a gang member, always a gang member"). Many databases have purging requirements (e.g., after 5 years of inactivity), which can be surmounted easily by entering new street contacts (field investigation cards) or even observations of associations with documented gang members.[2] It almost goes without saying that if such problems as these are found in local gang databases, then achieving the valued and reliable state databases (to say nothing of federal counterparts) urged by Barrows and Huff could be impossible.

1. These offenses are listed in LAPD's Citywide Gang Crime Summary, December 2008: Aggravated assault, robbery, homicide, shots at inhabited dwellings, criminal threats, hijackings, attacks on police officers, kidnap, extortion, rape, and arson.
2. In this regard, note that Barrows and Huff's (2009) report that Minnesota's purging controls can allow reinstating a person into the gang database regardless of whether a new conviction or adjudication involves a gang-related offense.

For the Defense

As a card-carrying member of the American Civil Liberties Union (ACLU), I was given an ACLU cap. In turn, after a friendly argument with a politically conservative neighbor, I presented the cap to him; he refers to it as his UCLA cap! As a frequent expert witness for the defense in gang cases, I have become nervous about how the prosecution, often using local gang databases, brings inappropriate definitions, one-sided stereotypes, and overstatements of gang identification of defendants to influence juries. Indeed, in the absence of the rare gang expert for the defense to battle the gang cop serving as expert witness for the prosecution, protection of civil rights becomes a real issue. It is not that gang defendants are innocent—in most of my cases, they seem not to be so—but that their punishments can be excessively severe because of implications of gang membership or gang-related intentions. Gang databases are seldom, if ever, an advantage to the defendant. If he (or the rare she) is not in the database, then the prosecution can elicit testimony as to the unreliability of the database. Attempts by the defense to use the unreliability argument are generally overcome by the gang cop's personal knowledge of the defendant ("I've seen him in the company of gang members" or "Other officers tell me he's gang affiliated"). In California and some other states, conviction for a "gang-related" offense can double the length of one's sentence, and misuse of the gang database easily contributes to this.

Databases and Program Evaluations

U.S. cities are replete with both public and private programs to prevent, intervene in, or suppress gang activity. Most programs are not evaluated for effectiveness, which leads to their continuation regardless of demonstrable outcome. In the few notable exceptions, evaluations have tended to show little or no positive effect and, occasionally, a negative outcome of increased gang size, cohesiveness, or illegal activity. As this essay was being written, a massive new gang prevention and intervention program was being initiated by the Los Angeles Mayor's Office, amounting to some $20 million every year for at least 4 years. Along with this program is an independent evaluation, budgeted at $900,000 every year.

And, as this essay was being written, the independent evaluation organization and the mayor's office were in serious discussions about appropriate outcome measures: Should it be LAPD arrest data or self-reported crime data? Strong convenience arguments exist for the LAPD data. Strong, if less convenient, arguments are present for the self-report approach. Those individuals arguing for both, to date, have not yet provided convincing arguments about how to meld the two. As one of several criminological advisors to the program, I have urged strongly against reliance on the LAPD data for all the reasons noted in this essay. We have not yet prevailed, but the discussion is ongoing. Some consideration exists of obtaining funding to develop a self-report crime database. Now wouldn't that be something!

Knowledge Building

Where do I come down? After many years of poking around in and of using local gang databases—police, probation, and court—I have become convinced that the administrative needs of these data-gathering organizations would work effectively against valid results and inconsistently for and against reliable results. I do not trust their databases except under certain circumstances in which they are used to compare cohorts equally exposed to the same data-gathering procedures in the same jurisdiction. For example, one could compare Gang A against Gang B, fringe versus core members, and a gang's activity during a period of time when the data-gathering procedures are stable. That is, I see some prospects for reliability, even if few for validity. And, as to Barrows and Huff's (2009) desire for building state-level gang databases, I find I have little sympathy so long as they are based on the aggregation of many local efforts. Questionable reliability, low validity, and intersource "noise" can lead to inappropriate policy and unacceptable criminological research.

References

Barrows, Julie and C. Ronald Huff. 2009. Gangs and public policy: Constructing and deconstructing gang databases. *Criminology & Public Policy*. This issue.

Malcolm W. Klein is professor emeritus of sociology at the University of Southern California. Aside from continued writing on street gang issues, his principal activities now include (1) involvement in the 12-year-old Eurogang Program, (2) expert witness testimony in scores of criminal gang cases, and (3) policy and research guidance for the multimillion-dollar gang prevention and intervention program for the Los Angeles Mayor's Office.

Gangs, law enforcement, and the academy

James F. Short, Jr.
Washington State University

Gangs have been a popular topic for *Criminology & Public Policy* (CPP) beginning with its initial issue (Esbensen, Osgood, Taylor, Peterson, and Freng, 2001). More recent forays have appeared on such topics as the use of civil gang injunctions (Grogger, 2005; Maxson, Hennigan, and Slone, 2005) and network analysis (McGloin, 2005; Papachristos, 2005).[1] Scott Decker (2007) argued for expansion and use of police gang units, and most recently, Malcolm Klein's 2009 Vollmer Award Address described his long career of balancing knowledge and policy concerns while studying gangs and other important topics (and was followed by commentaries from Scott H. Decker [2009] and Delbert S. Elliott [2009]). Now comes "Gangs and public policy: Constructing and deconstructing gang databases" by Barrows and Huff (2009, this issue).

At the most fundamental level, the problem with gang databases is that they are, *perhaps of necessity*, based on law-enforcement criteria with all that it entails. Why the caveat? Because it is possible, and I believe necessary, to conceive of databases that are more inclusive and broadly based than reports by police or other law-enforcement agencies, databases that involve other institutions and other levels of expertise (see Short and Hughes, in press).[2]

---

1. Tracy Meares edited and introduced the civil gang injunction and network exchanges. Papachristos (2005) compared "Deciphering and doing something about modern street gangs" with "Interpreting inkblots."

2. Short and Hughes (in press) argued that, in the interest of research integrity, evaluation of the efficacy of gang-control programs requires theoretically defensible criteria in addition to arrests and other involvement in criminal and juvenile justice systems. We advanced as candidates for such criteria evaluation of gang program efforts toward improving the social capital of offenders and collective efficacy of communities. In principle, the argument applies equally for local gang databases; i.e., gang and gang-member criteria should include identification of gang members and gangs based on multiple criteria that include, but not exclusively, police reports. For an excellent review of street gang programs and policies, see Klein and Maxson (2006).

That, indeed, was the hope of *The Little Village Gang Project in Chicago*, the large-scale gang violence reduction project directed by Irving Spergel (2007; Spergel, Wa, and Sosa, 2006). This project, and later attempted replications of the model on which it was based, documented many of the problems that inevitably develop when law enforcement and other agencies and institutions try to work together for program and research purposes. Although the Chicago Police Department (CPD) was the sponsor and lead agency of this ambitious collaborative project, a series of political and agency changes quickly required project reorganization "almost before it began" (Spergel, 2007: 33).

Spergel's (2007) description of these organizational and other problems is a reminder of how difficult it has been to achieve cooperation—let alone collaboration—for programs and research that attempt to bring together law enforcement and other agency and institutional programs and research, including independent program evaluations. Elaborate plans and high hopes often unravel with implementation. After more than 5 years and dashed hopes for continuation and a degree of permanency, the Little Village Project was terminated. Claims of success in reducing gang violence, documented by careful statistical and qualitative measures, were modest, as was the case with other "comprehensive, community-wide gang programs" modeled after the Little Village Project (Spergel et al., 2006).

All of these programs were marked by often heroic efforts, especially by Spergel and his team, to overcome the litany of problems posed by specialized law-enforcement units (e.g., police gang units) that Decker (2007: 732) warned against: "isolation, a lack of information sharing, an inability to penetrate community environments, the lack of links to other enforcement and prosecution agencies, and ... creation of conflict within the police organization and with the community."

Such problems notwithstanding, Spergel (2007: 342) concluded that effective control of gang crime, "particularly gang violence" requires that "law enforcement—under the aegis and control of a central unit of government, the mayor's office, or the city council ... be the lead agency, supported and influenced by an interagency and community advisory council, and a street-outreach team of police, probation, youth workers (including former gang influentials from gang-problem areas), neighborhood organizers, and associated school, treatment, and job-placement personnel collaboratively serving, controlling, and targeting problem gang youth."[3]

Similarly, Decker's (2007: 731) prescription for correcting the problems he identified is the expansion of police gang units that include "a strong focus on community policing"—a recommendation he acknowledged might seem suspect in view of evidence that "police gang units generally engage in behavior that is the antithesis of the principles of effective community policing" (Katz and Webb, 2006).

---

3.   In response to an earlier draft of this essay, Spergel commented via e-mail that the Little Village Project experienced problems with police primarily at the departmental and mayoral levels and that, contrary to the views and experience of many gang experts, the local district and gang-unit levels were cooperative and "committed to solving" gang problems.

Problems of gang-control programs such as those identified by Spergel (2007) and Decker (2007) add to a host of problems that compromise the reliability and the validity of much police-generated data on gangs. Decker and Spergel certainly were correct in stressing the importance of involving both the police and community institutions for successful gang control (see Klein and Maxson, 2006). At the most fundamental level, however, problems will remain as long as the police are the primary arbiters of gang and gang-member identity.

Decker (2007) noted that police departments have made huge strides in tactical gang intelligence in recent years. As Barrows and Huff (2009) explained, however, such information has led to Type 1 and Type 2 errors when identifying gangs and gang members. These errors in turn have resulted in abuses related to race and ethnicity and in—sometimes gross—mis-representation and misunderstanding of the nature of gangs, their behavior, and their location within community contexts.

Addressing the latter of these problems first, Mercer Sullivan's (2006) study of information sources regarding gangs and crime is instructive. Sullivan compared media reports, police arrests, and field-work data in three New York City (NYC) neighborhoods during the 1990s—a period of gang reemergence in the city. Historically, gang phenomena in NYC have followed cyclical patterns, often peaking in response to particular events. Sullivan documented peaks and declines for media reports of NYC gang activity during the decade, although violent felony arrests were generally declining in the city as well as in the three neighborhood areas where field observations and interviews with young people were conducted. Although gangs were not the primary focus of the field studies, "the topic of gangs surfaced repeatedly" and "the meanings of the term *gang* were contested throughout" (p. 25, emphasis in original). Comparison of the three neighborhoods revealed similarities and differences, including "rapid shifts, even within a single ecological area, in how youths enact violence and *what they call it* when they do" (p. 32, emphasis added). Among the observed similarities were "the transience and ambiguity of gang labels" (p. 33). Sullivan (p. 34) concluded by warning against *reification* of gangs:

> The problem with gang studies as a distinct category of criminological inquiry is that *gang studies*, by virtue of the very title, implies answers to questions that should remain open to inquiry. . . . Generic gang prevention programs that concentrate on the outward symbols of group membership at the expense of concentrating on the specific local contexts of youth violence can shortchange the development of effective, knowledge-based, problem-oriented solutions. Reifying youth gangs imperils the study of youth violence.[4]

Other close-to-the-street studies have confirmed the limitations of law-enforcement data for understanding gangs and gang members (for insightful reviews, see Coughlin and Venkatesh, 2003; Klein and Maxson, 2006; Vigil, 2003). As the Barrows and Huff (2009) article notes,

---

4. Sullivan here echoed Zimring and Hawkins (1997).

it is precisely the "outward symbols" and number of "gang members," in addition to varying combinations of criminal activities, upon which most state laws focus.

No state law specifies racial or ethnic identity in defining gangs and gang members, of course, but substantial evidence exists that police have done so in some cities (Leyton, 2003; Zatz and Krecker, 2003). And although racial profiling studies have not focused specifically on gangs, the consequences of identifying individuals and small groups of minority youth as gang members (individually or collectively) are likely to be especially serious and may be fateful compared with "stop and search/frisk" practices that research indicates are inefficient, ineffective, and irresponsible and, more importantly, that undermine trust in the law and law enforcement (Alpert, 2007).

In ways that are not altogether clear, these problems reverberate in local communities and throughout the juvenile and criminal justice systems. Reification of gangs occurs at every level: among local community residents and the general public, legal institutions, academic research, and gang members.[5] The resulting ambiguity strongly cautions against direct social policy application of any gang definition or identification without careful, systematic attention to local community contexts. Without such attention, statutory policies and practices attributing increased seriousness of criminal behavior to gang membership are especially consequential in that they distort the nature of gangs and their behavior, and they are likely to weaken and perhaps be counterproductive with respect to crime control.

The blurring of distinctions between juvenile- and adult-justice systems—the subject of much research and commentary (see Short and Hughes, 2008)—adds to the importance of this issue. Barry Krisberg's (2003) chapter, "The end of the juvenile court: Prospects for our children," is apposite. Krisberg was especially concerned with the trend of increasingly transferring juveniles to criminal courts—despite the lack of solid evidence that transfers are a deterrent to criminal behavior or are constructive in other ways. Adding gang membership to this trend confounds an already complex and confusing mix of laws and policies affecting juveniles, especially racial and ethnic minorities.

Problems such as these might be dismissed as expected "growing pains" toward the achievement of comprehensive, collaborative programs and research efforts. Data systems do, after all, evolve, as evidenced by continuing evaluations and changes in such systems as the Uniform

---

5.   A study of Arizona anti-gang laws and their enforcement is illustrative. Marjorie Zatz and Richard Krecker (2003) noted that Arizona law adds a point to risk assessment of juveniles for suspected gang membership as well as criminalizing gang membership, enhancing sentence severity, and permitting police to testify in court as expert witnesses regarding gang membership. "As a set these laws increase the penalties for offenses committed by individuals who have been identified as gang members, solely because of that identification" (Zatz and Krecker, 2003: 179, emphasis in original). The result is that Latinos and Latinas are overrepresented in gang statistics, as both they and Blacks are in national crime data. Their interviews with Maricopa County juvenile court officials and defense attorneys found that the latter often challenge these laws on the grounds that "allegations of gang membership" are fortuitous (p. 183) and that court officials tend to regard allegations of gang affiliation as "hearsay" (p. 184). However, judges now have less latitude in handling juveniles than was the case prior to the statutory changes.

Crime Reports (UCR), which include innovations such as the Supplementary Homicide Reports (SHR) and the National Incident Based Reporting System (NIBRS)—both of which have been subjected to continuous commentary and evaluation. The National Crime Victims Survey (NCVS), under the aegis of the Bureau of Justice Statistics, recently has been the subject of intense evaluation by committees of the National Research Council.[6]

Both the SHR and NIBRS distinguish between "Gangland killings" and "Juvenile gang" involvement in homicides (SHR) and "Aggravated assault/homicide" (NIBRS); NIBRS also codes for "circumstances" of aggravated assault/homicide. The NIBRS manual, recently updated (2009: 22) defines a gang as "an ongoing organization, association or group of three or more persons, have a common interest and/or activity characterized by the commission of or involvement in a pattern of criminal or delinquent conduct" and distinguishes between "juvenile," "other," and none/unknown." The NCVS asks respondents whether single or multiple offenders were member(s) of a street gang. These systems do not identify individual offenders or victims, and they do not, as a matter of policy (or practice, generally), target minority populations or the poor. Years of data accumulation in these systems, particularly regarding serious crimes of violence, permit evaluation of their reliability, and to some extent, their validity.

Gang specialists to whom I sent an earlier draft of this essay made the point that police databases have been useful for investigative purposes, which is a major police function. The data have also become better as a result of challenges in court. Such challenges could strengthen both their reliability and their validity for police purposes (see Katz, Webb, and Schaefer, 2000). Additional research focused on local contexts would enhance their usefulness for scientific purposes as well, but by definition, incorporating such information into state or national gang databases would not be feasible.

The Division of Labor as a Solution

The centrality of the police to crime control is not in question and is much evidenced in Todd Clear and Natasha Frost's (November 2007) final "gala" CPP issue. Many of the more than two dozen policy discussions in that issue directly involved the police; indirectly, virtually all did.[7]

The relatively straightforward distinctions made by the SHR, NIBRS, and the NCVS are useful for monitoring serious crimes of violence on a national basis, particularly if programs are improved to take advantage of "twenty-first century information technologies and analysis capabilities," and the UCR is transferred "from the FBI to the Bureau of Justice Statistics," as Richard Rosenfeld (2007: 830) suggested.

---

6.   With few exceptions, criminologists have made little use of gang data from these sources.

7.   Space limitations preclude discussing the implications for gang databases of the emergence in late modernity of "the new regulatory state" (Braithwaite, 2000). Suffice to say that, with the privatization of many police functions, the study of regulatory policies and practices is likely to become of greater importance to criminology, including studying youth groups. In any case, "the key public policy nut to crack," as Braithwaite argued, is likely to remain "the ordering of local knowledge" (p. 65).

More caveats are necessary, however. The often counterproductive consequences of gang suppression and "zero tolerance" policies favored by many police departments, as well as local and national governments, is well documented (Hagedorn, 2008; Zilberg, 2007). The consequences of official definitions of gangs and gang members for those identified as such can be fateful, as social constructionists from Tannenbaum to Becker and from Bourdieu to Foucault have argued. Adding civil rights violations to labeling and constructionist issues continues to weaken proposals for state or national gang databases.[8]

The police should not be expected to perform systematic studies of youth collectivities, whether they are called gangs or something else. Such research is the job of scholars, independent of policy burdens. Imaginative approaches toward that end continue to be made (Decker and Weerman, 2005; Hagedorn, 2008; Short and Hughes, 2006). Nor should police be expected to understand gangs and gang members in precisely the terms that scholarly work employs, but their identification of gangs and gang members should be informed by these studies, and their behavior toward gangs and gang members should be informed by scholarly work in coordination with their constituent communities.

In historical perspective, controversy concerning the conceptual status of gangs is akin to that which followed Edwin Sutherland's introduction of the notion of white-collar crime. Was it really crime, given that phenomena so classified were not always in violation of criminal statutes? Were violators of organizational regulations really criminals, because many did not regard their behaviors as crime or consider themselves to be criminals? Research and theory on white-collar crime have moved beyond these controversies as scholars have studied different types of deviant—if not always criminal—behavior by individuals, organizations, and institutions (cf. Vaughan, 1983, 1996).

The study of youth collectivities such as gangs, however defined, has yielded a rich body of knowledge that is neither bound to nor bounded by official definitions. Identification of specific networks, for example, which include those based on police data, especially for more serious offenses that are more likely to be reported (one of the original criteria for development of the UCR), are enlightening and potentially useful, particularly when combined with field research (e.g., Papachristos, 2009).

Controversy concerning the measurement and the realities of law making and breaking, delinquency and crime, delinquents, criminals, and gangs will never cease. *Indicators used for social decision making are always subject to distortion of the phenomena they represent and to corruption of*

---

8. Conflation of street gangs and terrorists, which is both counterfactual and counterproductive but suggested by the U. S. Department of Homeland Security, raises more civil rights issues (Hagedorn, 2008). A National Gang Crime Research Center study (Knox, 2005) argued that the recruitment of imprisoned gang members into terror networks is a serious problem in the United States. In response to the somewhat leading question, "In your opinion, would an official who tries to bargain with an inmate gang leader be similar to negotiating with terrorists?", more than 70% of surveyed prison security threat group coordinators agreed.

the social processes they are intended to monitor.[9] Law enforcement and the academy need quantitative indicators, but should also be skeptical of claims that, because a measure is quantitative, it is therefore objective. Certainly, this statement is true of crime and gang statistics.

References

Alpert, Geoffrey P. 2007. Eliminate race as the only reason for police–citizen encounters. *Criminology & Public Policy*, 6: 671–678.

Barrows, Julie C. and C. Ronald Huff. 2009. Gangs and public policy: Constructing and deconstructing gang databases. *Criminology & Public Policy*. This issue.

Braithwaite, John. 2000. The new regulatory state and the transformation of criminology. In (David Garland and Richard Sparks, eds.), *Criminology and social theory*. Oxford, UK: Oxford University Press.

Clear, Todd R. and Natasha A. Frost. 2007. Informing public policy. *Criminology & Public Policy*, 6: 633–640.

Coughlin, Brenda C. and Sudhir Alladi Venkatesh. 2003. The urban street gang after 1970. *Annual Review of Sociology*, 29: 41–64.

Decker, Scott H. 2007. Expand the use of police gang units. *Criminology & Public Policy*, 6: 729–734.

Decker, Scott H. 2009. An essay in tribute to Malcolm Klein on his recognition as the August Vollmer Award Winner. *Criminology & Public Policy*, 1: 13–18.

Decker, Scott H. and Frank M. Weerman (eds.). 2005. *European street gangs and troublesome youth groups*. Lanham, MD: AltaMira.

Elliott, Delbert S. 2009. Malcolm W. Klein: August Vollmer Award Winner, 2008. *Criminology & Public Policy*, 1: 19–23.

Federal Bureau of Investigation. 2009. *Developments in the national incident-based reporting system (NIBRS)*. Washington DC: U.S. Department of Justice.

Esbensen, Finn-Aage, D. Wayne Osgood, Terrance J. Taylor, Dana Peterson, and Adrienne Freng. 2001. How great is G.R.E.A.T? *Criminology & Public Policy*, 1: 87–118.

Grogger, Jeffrey. 2005. What we know about gang injunctions. *Criminology & Public Policy*, 4: 637–642.

Hagedorn, John M. 2008. *World of gangs: Armed young men and gangsta culture*. Minneapolis: University of Minnesota Press.

Katz, Charles M. and Vincent J. Webb. 2006. *Policing gangs in America*. New York: Cambridge University Press.

Katz, Charles M., Vincent J. Webb, and David R. Schaefer. 2000. The validity of police gang intelligence lists: Examining differences in delinquency between documented gang members and non-documented delinquent youth. *Police Quarterly*, 3: 413–437.

Klein, Malcolm W. 2009. Bootlegging: A career caught between fantasy and reality—Vollmer Award Address. *Criminology & Public Policy*, 8: 1–12.

9. Somewhere in my files, I have a quotation to this effect attributed to Donald Campbell I believe, but I cannot be certain.

Klein, Malcolm W. and Cheryl L. Maxson. 2006. *Street gang patterns and policies*. New York: Oxford University Press.

Knox, George W. 2005. *The problem of gangs and security threat groups (STGs) in American prisons today: Recent research findings from the 2004 Prison Gang Survey*. Retrieved June 29, 2009 from ngcrc.com/corr2006.html.

Krisberg, Barry A. 2003. The end of the juvenile court: Prospects for our children. In (Darnell Hawkins, Samuel L. Myers, and Randolph F. Stone, eds.), *Crime control and social justice: The delicate balance*. Westport, CN: Greenwood Press.

Leyton, Stacey. 2003. The new blacklist: The threat to civil liberties posed by gang databases. In (Darnell Hawkins, Samuel L. Myers, and Randolph F. Stone, eds.), *Crime control and social justice: The delicate balance*. Westport, CN: Greenwood Press.

Maxson, Cheryl L., Karen M. Hennigan, and David C. Sloane. 2005. It's getting crazy out there: Can a civil gang injunction change a community? *Criminology & Public Policy*, 4: 577–606.

McGloin, Jean Marie 2005. Policy and intervention considerations of a network analysis of street gangs. *Criminology & Public Policy*, 4: 607–636.

Papachristos, Andrew V. 2005. Interpreting inkblots: Deciphering and doing something about modern street gangs. *Criminology & Public Policy*, 4: 643–652.

Papachristos, Andrew V. 2009. Murder by structure: Dominance relations and the social structure of gang homicide. *American Journal of Sociology*. In press.

Rosenfeld, Richard. 2007. Transfer the uniform crime reporting program from the FBI to the Bureau of Justice Statistics. *Criminology & Public Policy*, 6: 825–833.

Short, James F., Jr. and Lorine A. Hughes (eds.). 2006. *Studying youth gangs*. Lanham, MD: AltaMira Press.

Short, James F., Jr. and Lorine A. Hughes. 2008. *Juvenile delinquency and delinquents: The nexus of social change*. Upper Saddle River, NJ: Prentice Hall.

Short, James F., Jr. and Lorine A. Hughes. In press. Promoting research integrity in community-based intervention research. In (Robert Chaskin, ed.), *Youth gangs and community intervention: Research, practice, and evidence*. New York: Columbia University Press.

Spergel, Irving A. 2007. *Reducing gang violence: The Little Village Project in Chicago*. Lanham, MD: AltaMira.

Spergel, Irving A., Kwai Ming Wa, and Rolando Villarreal Sosa. 2006. The comprehensive, community-wide gang program model: Success and failure. In (James F. Short, Jr. and Lorine A. Hughes, eds.), *Studying youth gangs*. Lanham, MD: AltaMira.

Sullivan, Mercer L. 2006. Are "gang" studies dangerous? Youth violence, local context, and the problem of reification. In (James F. Short, Jr. and Lorine A. Hughes, eds.), *Studying youth gangs*. Lanham, MD: AltaMira.

Vaughan, Diane. 1983. *Controlling unlawful organizational behavior: Social structure and corporate misconduct*. Chicago, IL: University of Chicago Press.

Vaughan, Diane. 1996. *The Challenger launch decision: Risky technology, culture, and deviance at NASA*. Chicago, IL: University of Chicago Press.

Vigil, James Diego. 2003. Urban violence and street gangs. *Annual Review of Anthropology*, 32: 225–242.

Zatz, Marjorie S. and Richard P. Krecker, Jr. 2003. Anti-gang initiatives as racialized policy. In (Darnell Hawkins, Samuel L. Myers, and Randolph F. Stone, eds.), *Crime control and social justice: The delicate balance*. Westport, CN: Greenwood Press.

Zilberg, Elana. 2007. Refugee gang youth: Zero tolerance and the security state in contemporary U.S.–Salvadoran relations. In (Sudhir Alladi Venkatesh and Ronald Kassimir, eds.), *Youth, globalization, and the law*. Stanford, CT: Stanford University Press.

Zimring, Franklin E. and Gordon J. Hawkins. 1997. *Crime is not the problem: Lethal violence in America*. New York: Oxford University Press.

James F. Short, Jr. is professor emeritus of sociology at Washington State University. The author of numerous books, including *Delinquency, crime, and society* (University of Chicago Press, 1976) and *Poverty, ethnicity, and violent crime* (Westview Press, 1997), Short also served as co-director of research for President Lyndon Johnson's Commission on the Causes and Prevention of Violence and as a member of the advisory board for the Social Science Research Council Task Force on Hurricane Katrina and Rebuilding the Gulf Coast. Short is a past president of the American Sociological Association and the American Society of Criminology. He has also served as editor for the American Sociological Review. Awards and honors include a John Simon Guggenheim Fellowship, a National Institute of Mental Health Fellowship, the Edwin H. Sutherland Award (American Society of Criminology), the Bruce Smith Award (Academy of Criminal Justice Sciences) and the American Sociological Association's Award for Distinguished Contribution. Short is also a fellow of the American Association for the Advancement of Science.

P O L I C E   M I S C O N D U C T

Conceptual, methodological, and policy considerations in the study of police misconduct

James Frank, *Senior Editor*
University of Cincinnati

S ince the formation of formal police departments, police misconduct and the abuse of police powers has been a concern (Lundman, 1980). Unfortunately, police misconduct has often manifested itself in the media in the form of publicized scandals resulting in the formation of high-profile political commissions (i.e., Wickersham, Kerner, and Mollen) and in the dismissal of police personnel. Because police are the gatekeepers of social control, reports of power abuse have created tension between the police and the public. The failure to control officer discretion and misconduct results in crises for police organizations (Mastrofski, 2004). Throughout time, police misbehavior has received considerable attention from police researchers, police administrators, and policymakers. Academic discourse has often concerned the extent of misconduct, the nature of misconduct, the extent to which misconduct permeates police organizations, and the preferred strategy for controlling officers.

This topic is extremely important today because police agencies continue to struggle with issues related to police accountability, as well as with the selection, hiring, and retention of quality police officers. The presumption is that the selection of quality personnel translates into effective crime fighting, positive community interaction, and improvements in police accountability (Sanders, 2003; White, 2008). As Sanders contends, it is much easier to recognize and measure bad performance than good performance. Although consensus is lacking as to what is quality officer performance, it is recognized that nonfeasance, excessive use of force, discourteous and aggressive actions, and other forms of police misconduct represent traits of bad officers.

The article by Kane and White (2009, this issue) examines police misconduct resulting in an officer being formally dismissed by the police organization. Using data collected from personnel files of 1,543 New York Police Department (NYPD) officers hired between 1975 and

Direct correspondence to James Frank, University of Cincinnati, Division of Criminal Justice, P.O. Box 210389, Cincinnati, OH 45221-0389 (e-mail: James.Frank@uc.edu).

1996 who were formally dismissed by the agency for police misconduct—and a comparison sample of officers selected from the same hiring cohort who served honorably—Kane and White examine the correlates of formal department dismissals. Their findings highlight the influence of risk and supportive factors related to the formal detection and processing of misconduct as well as variation in the influence of certain characteristics on three categories of misconduct. Because police misconduct involves the actions of officers, it is not surprising that individual-level characteristics correlate with misbehavior (Travis and Langworthy, 2008). Their determination that prior criminal involvement, early employment problems with the police agency, history of citizen complaints, and education level significantly predict formal dismissals, however, is extremely important for police policymakers involved in devising evidence-based selection and hiring policies, implementing early warning systems (Hughes and Andre, 2007; Walker, Alpert, and Kenney, 2000), and improving police accountability.

Kane and White's (2009) article also highlights the difficulty accounting for findings that involve the influence of officer race when predicting officer behavior (Brown and Frank, 2006; Sun and Payne, 2004; Sun, Payne, and Wu, 2007). One primary reform effort in law enforcement from the past 40 years has been to diversify police forces by increasing the number of minorities in policing. Despite increases in their representation (11% to 13 % depending on the data source), minorities remain underrepresented in small and medium departments as well as in top-level administrative positions (National Research Council, 2004). As Kane and White admit—and as the policy essays by Manning (2009, this issue), King (2009, this issue) and Ivkovich (2009, this issue) highlight—several interpretations exist for the findings that African Americans were dismissed at a higher rate than Caucasian officers and that officer race was related to two categories of misconduct. Organizational involvement in the detection and processing of misconduct could result in differential processing of misconduct by minority officers. Similarly, group loyalty associated with the police culture might not extend to minority officers, which makes it easier for agencies to detect and process the misconduct. Finally, minority officers could be involved in higher levels of certain forms of misconduct. Each of these explanations might be plausible.

The policy essays by Manning (2009), King (2009) and Ivkovich (2009) highlight obstacles to conducting research on police misconduct, namely the preferred methodological approach for studying this phenomenon, the extent and nature of officer misconduct, as well as the level of organization within police agencies, and alternatives for controlling police misbehavior. First, determining the extent of police misconduct is an almost impossible task. Because of the low visibility of street-level officer behavior, instances of police misconduct do not come to the attention of most observers. Informal rules of the police culture inhibit peers from reporting instances of misconduct and increase the difficulty of discovering misconduct (Haar, 1997; Haar and Morash, 1999). Furthermore, disagreements surround the definition of misconduct and the included behaviors. Other than the most egregious behaviors that include law violations, debate persists about the inclusion of behaviors on the lower end of the misconduct continuum, and

definitions may vary by agency. Finally, similar to the use of force, police misconduct is a rare event. Irrespective of the research methodology employed (e.g., ethnographic, official reports), we know that most police officers do not engage in work-related misconduct. Kane and White (2009), however, construct a dataset that includes a sufficient number of formally processed misconduct cases to allow for the examination of dismissals for cause.

Second, similar to other phenomenon, data collection concerning acts of misconduct is difficult, and the use of official data has certain limitations. Admittedly, the use of official data is likely to undercount misconduct occurrences. For misconduct to become recorded in official records, it must first be detected and then be formally processed. This is complicated by Manning's (2009) observation that organizations might not want some forms of misconduct to appear in department reports and agency files might not contain all relevant information.

Third, police officers work within an organizational, political, and social context, and they serve multiple audiences. As such, officer behavior is difficult to control, although organizational strategies have been effective at controlling officer behavior (Mastrofski, 2004). King (2009), Ivkovich (2009), and Manning (2009) emphasize the importance of the organization in controlling officer behavior. King discusses HROs (highly reliable organizations) and changing the norms and values of the organizational culture. Ivkovich addresses the need to influence the police culture and the need to examine suborganizational units (i.e., precincts and bureaus) for variation in the commission, detection, and processing of misconduct. Furthermore, Manning suggests a cohort analysis of misconduct. Ultimately, it is the responsibility of the police organization to control the exercise of officer discretion.

Finally, Manning (2009), King (2009), and Ivkovich (2009) suggest directions for future research that likely would contribute to our understanding of the nature and extent of police misconduct as well as agency responses to police abuse of power. As noted by Manning and Ivkovich, a precinct-level analysis could highlight organizational differences in both the extent and the sanctioning of misconduct. Differences might occur in oversight as to the tolerance of misconduct and in the willingness to engage the organization in formally processing offenders. A focus on the organizational context could provide information on the detection and processing of misconduct by minority officers.

References

Brown, Robert and James Frank. 2006. Race and officer decision making: Examining differences in arrest outcomes between Black and White officers. *Justice Quarterly*, 23: 96–126.

Haar, Robin. 1997. Patterns of interaction in a police patrol bureau: Race and gender barriers to integration. *Justice Quarterly*, 14: 53–58.

Haar, Robin and Merry Morash. 1999. Gender, race, and strategies of coping with organizational stress in policing. *Justice Quarterly*, 16: 303–36.

Hughes, Frank V. and Lisa Andre. 2007. Problem-officer variables and early-warning systems. *The Police Chief*, 74: 164–72.

Ivkovich, Sanja Kutnjak. 2009. Rotten apples, rotten branches, and rotten orchards: A cautionary tale of police misconduct. *Criminology & Public Policy*. This issue.

Kane, Robert J. and Michael D. White. 2009. Bad cops: A study of career-ending misconduct among New York City police officers. *Criminology & Public Policy*. This issue.

King, William R. 2009. Police officer misconduct as normal accidents: An organizational perspective. *Criminology & Public Policy*. This issue.

Manning, Peter K. 2009. Bad cops. *Criminology & Public Policy*. This issue.

Lundman, Richard. 1980. *Police and policing: An introduction*. New York: Holt, Rinehart and Winston.

Mastrofski, Stephen. 2004. Controlling street-level police discretion. *Annals*, 593: 100–118.

National Research Council. 2004. *Fairness and effectiveness in policing: The evidence committee to review research on police policy and practices* (Wesley G. Skogan and Kathleen Frydl, eds.). Committee on Law and Justice, Division of Behavioral and Social Sciences and Education. Washington, DC: National Academies Press.

Sanders, Beth. 2003. Maybe there's no such thing as a "good cop": Organizational challenges in selecting quality officers. *Policing: An International Journal of Police Strategies and Management*, 26: 313–328.

Sun, Ivan and Brian Payne. 2004. Racial differences in resolving conflicts: A comparison between Black and White police officers. *Crime & Delinquency*, 50: 516–541.

Sun, Ivan, Brian Payne, and Yenning Wu. 2007. The impact of situational factors, officer characteristics and neighborhood context on police behavior: A multilevel analysis. *Journal of Criminal Justice*, 36: 22–32.

Travis, Lawrence and Robert Langworthy. 2008. *Policing American: A balance of forces*, 4th Edition. Upper Saddle, River, NJ: Pearson/Prentice Hall.

Walker, Samuel, Geoff Alpert, and David Kenney. 2000. Early warning systems for police: Concept, history and issues. *Police Quarterly*, 3: 132–152.

White, Michael. 2008. Identifying good cops early: Predicting recruit performance in the academy. *Police Quarterly*, 11: 27–49.

James Frank is a professor in the Division of Criminal Justice and a director of the Center for Criminal Justice Research at the University of Cincinnati. He received his Ph.D. from the School of Criminal Justice at Michigan State University and his J.D. from Ohio Northern University. His primary research interests include understanding police behavior at the street level, the formation of citizen attitudes toward the police, and the use of evolving technology by patrol officers. He has published policing articles in *Justice Quarterly*, *Police Quarterly*, *Journal of Criminal Justice,* the *American Journal of Police*, *Crime & Delinquency,* and *Policing: An International Journal of Police Strategy and Management.*

Bad cops

A study of career-ending misconduct among New York City police officers

Robert J. Kane
Michael D. White
Arizona State University

Research Summary

Police scholars and public policymakers throughout generations have sought to identify reliable correlates of police misconduct. Despite these efforts, general statements as to the etiology and epidemiology of police misconduct remain inconclusive, in part because of the inconsistent definitions of misconduct and the difficulty of obtaining the data required to make such statements. This research attempts to fill these gaps through a comparison of the personal and career histories of all 1,543 officers who were involuntarily separated from the New York City Police Department (NYPD) for cause during 1975–1996 with a randomly selected sample of their police academy classmates who served honorably. The study uses confidential NYPD files as its major data sources, which include extensive bio-graphical and career information. The study finds that career-ending misconduct rarely occurs in the NYPD and that the types of misconduct do not match well with existing definitions. Several factors emerge as significant predictors of misconduct, including officer race, minimal education, records of prior criminality and prior poor employment, failure to advance in the NYPD, and histories of citizen complaints.

Policy Implications

This study shows that existing definitions of police misconduct are difficult to apply to actual cases of police malpractice, which leads the authors to create a new eight-category classification scheme. The rarity of misconduct, especially on-duty abuse, confirms prior research indicating that most police officers do their jobs without engaging in serious malpractice. These findings suggest that the NYPD has become better behaved as it has become more diverse along race and gender dimensions and that the link between black officers and misconduct might be explained by persistent "tokenism." The findings related to race have important implications for continued efforts to build racially representative police departments. Personal history findings highlight the importance of conducting background investigations that disqualify candidates with arrest records and employment disciplinary histories, whereas the inverse relationship between college education and misconduct provides strong support for continued emphasis on pre- and post-employment educational requirements.

Keywords

police misconduct, NYPD, scandal, tokenism, involuntary termination

P olice scholars and public policymakers throughout generations have sought to identify reliable correlates of police misconduct, focusing primarily on the organizational and individual units of analysis. Some researchers, for example, have examined links between personality (Christopher Commission, 1991; Muir, 1977; Toch, 1996) and static (Hickman, Piquero, Greene, 2000; Waugh, Ede, and Alley, 1988) characteristics and police brutality; others have tested associations between personal history factors (e.g., prior employment and military experience, previous arrests, and substance abuse) and general forms of misconduct (e.g., Cohen and Chaiken, 1972, 1973; Mollen Commission, 1994); and a few have investigated the effects of education on citizen complaints (Bowker, 1980; Cohen and Chaiken, 1972, 1973; Kappeler, Sapp, and Carter, 1992). Research findings on the correlates of police misconduct at the individual level have been largely mixed and are often ambiguous. As McManus (1969) noted, because police officers have long been chosen from a pool of people whose characteristics and backgrounds are homogenous, literature on the relationships between individual characteristics and deviance remains limited because of the lack of variability among samples. More recent work has confirmed these observations (Grant and Grant, 1996).

Research on police organizations generally has produced more reliable results, perhaps because organizational opportunity structures for misconduct are less idiosyncratic than the behaviors of individuals who might seek to exploit them. Skolnick and Fyfe (1993) argued that police administrators are often responsible for creating organizations that are tolerant of violence among officers as a way of maintaining street-level order and respect for authority (see also Chevigny, 1969, 1995). As such, officers who work in police agencies characterized by a so-called "siege mentality" (e.g., Skolnick and Fyfe, 1993: 106) can engage in excessive force

and other forms of abuse with the implicit protection and perhaps even encouragement of departmental management. Other research has identified discrete cycles of "scandal and reform" (e.g., Sherman, 1978), arguing that variations across time in corruption controls and the vigor with which agencies investigate allegations of misconduct are largely responsible for variations in organizational misbehavior, primarily in the form of organized corruption (see also Knapp Commission, 1972; Mollen Commission, 1994).

To date, general statements as to the etiology and epidemiology of police misconduct remain inconclusive, in part because of the inconsistent definitions of misconduct across studies and the difficulty of obtaining the data required to make such statements. Although some recent studies have increased the theoretical rigor with which hypotheses about police misconduct are evaluated (e.g., Chappell and Piquero, 2004; Kane, 2002), they are nevertheless limited in their scope and ability to support inferences about individual-level patterns of misconduct within the contexts of changing organizational structures.

The current study fills these gaps by reporting findings from a large-scale examination of police misconduct conducted in the nation's largest police department. Specifically, the study examines the life and career histories of all officers who were involuntarily separated from the New York City Police Department (NYPD) for reasons of misconduct (N = 1,543) from 1975 to 1996. These "study" officers were matched to a random sample of their police academy classmates who served without engaging in career-ending misconduct (N = 1,542), and the groups were compared across a broad range of personal history and employment variables. Notably, this examination of individual officers is anchored by two major public scandals that resulted in the Knapp Commission of 1972 and the Mollen Commission of 1994. The study addresses two broad questions related to police deviance: What was the nature and prevalence of organizational misconduct during this 22-year period, which included a fairly complete cycle of scandal and reform? And what were the relevant personal and career-history characteristics most responsible for distinguishing between officers who engaged in career-ending misconduct and officers who did not? The current study seeks to provide the academic and practitioner fields with evidence-based recommendations about the risk and protective factors associated with police misconduct.

Literature Review

Research has identified several categories of police deviance that inform the measurement of the current study's dependent variable. As a starting point for conceptualizing police misconduct, Kappeler, Sluder, and Alpert (1998) observed that police occupational deviance can be defined by police administrators through policies and procedures that govern officer conduct, as well as by the public—through criminal and civil statutes, Supreme Court decisions, citizen review boards, and so on. These "internal" and "external" typifications (Kappeler et al., 1998), however, are not always mutually exclusive: In many police departments, it is administratively impermissible (via policy) for its members of service to be arrested. Thus, if an officer who is

arrested is not subsequently convicted (i.e., case dismissed or found not guilty), then he or she still might be held accountable within the police organization for the initial arrest. As a result, even though police department policy and procedures manuals identify the boundaries of administrative conduct, they also sanction the most common forms of police misconduct that external stakeholders have defined as impermissible. These forms of misconduct include:

- *Police crime* occurs when officers use their positions of public trust to violate existing criminal statutes (Kappeler et al., 1998). Examples include stealing contraband from the police evidence lockers and extorting money from legitimate business owners for police protection.
- *Police corruption* involves officers who use their position or authority to engage in profit-motivated misconduct for personal gain (e.g., Sherman, 1978). Police taking money from illegitimate businesses in exchange for non-enforcement is an example of corruption.
- *Abuse of authority* includes physical abuse (e.g., police brutality and/or excessive force), psychological abuse (most often in the context of police interrogations), and legal abuse (generally in the form of perjury to achieve an organizational goal or to protect corrupt enterprises; see Carter, 1985; Kappeler et al., 1998).

A Brief Note on Off-Duty Misconduct

Not all police misconduct occurs when officers are on duty. Fyfe (1980a) reported that in New York City, 20% of police firearm discharges involved officers who were off duty and that the NYPD found cause for disciplinary or criminal action in half those cases. He subsequently reported a similar pattern of inappropriate off-duty police shootings—as well as lesser degrees of force—in Philadelphia (Fyfe, 1987, 1988). Moreover, off-duty officers can abuse their police authority while engaged in any number of other activities such as personal disputes, domestic violence, bar fights, drunk driving and related vehicle accidents, acts of vandalism, sex offenses, and larceny.[1]

Limitations of Existing Classifications of Police Misconduct

Existing classifications of police misconduct are limited in many ways. For example, Sherman (1978) noted that police crime does not describe all crimes committed by police officers because many offenses may have nothing to do with officers' employment status. Examples include off-duty burglaries, domestic assaults, or tax evasion—all of which certainly are crimes, but absent abuse of the *police* authority to gain the opportunity to commit the crimes, these acts of deviance should not be considered *police* crime. Also, police corruption is not always distinct

---

1. For example, several study officers were forced to leave the department as the result of abusing their authority at minor traffic accidents they caused while driving their personal vehicles. In such cases, officers would collide with another vehicle (usually on one of the bridges that connect Manhattan to the outer boroughs and often on the way to work), engage the other motorist in a verbal dispute, and when/if the other motorist engaged the officer verbally in the dispute, the officer would identify himself/herself as a New York City Police Officer, draw his/her off-duty weapon, and prone the other motorist out on the ground. In several of these instances, the officer shot the other motorist as part of the dispute process.

from police crime. Each category often includes some overlap with the other. Moreover, Hale (1989) noted that differing views exist of whether definitions of corruption should include only illegal behaviors. McMullan (1961), for example, noted that any public official is corrupt if he or she accepts compensation for not performing regular duties or for performing duties normally proscribed by their employment positions (which may or may not be illegal).

Last, although physical abuse is often considered a single construct, Fyfe (1986) distinguished between extralegal and unnecessary police violence. Fyfe argued that extralegal force, or brutality, represents intentional physical abuse inflicted maliciously and for no legitimate police purpose (see also Van Maanen, 1978). By contrast, unnecessary force typically results from police incompetence or carelessness, and it is generally not the product of malice. It usually occurs when officers unnecessarily put themselves in harm's way by using poor tactics while approaching potentially violent persons or situations. Then, when potential violence suddenly becomes real, the officers find that their exposed and vulnerable positions have left them no option but to resort to force to defend themselves (Fyfe, 1986).

Evidence on the Correlates of Police Misconduct
With the exception of a few recent studies (e.g., Chappell and Piquero, 2004; Kane, 2002), most identified research on police misconduct has been conducted in the absence of rigorous theoretical frameworks. This might be partially because of the difficulty of fully defining the police deviance construct. Although a substantial portion of police misconduct represents administrative nonconformity (similar to nonconformity in other occupational settings) and fits well with organizational theories, a great deal of police misconduct is illegal and can be explained by traditional criminological or justice theories. Furthermore, and as noted, with respect to police deviance—perhaps more than other occupations—it is often difficult to distinguish administrative nonconformity from violations of the law: Extorting money from drug dealers in exchange for not arresting them is both illegal and administratively impermissible. Thus, criminological theories that might explain criminal police misconduct could also explain some patterns of administrative deviance among officers.

Chappell and Piquero (2004) used social learning theory to distinguish between officers who received complaints and those who did not in a sample of Philadelphia police officers. The researchers operationalized misconduct as theft (illegal), excessive use of force (likely illegal), and accepting gifts and free meals from members of the public (administratively impermissible but not illegal). Although Chappell and Piquero engaged in a creative and useful application of social learning theory, their indicators of misconduct were derived from scenario-based vignettes of hypothetical situations. The researchers did not study actual misconduct events.

Kappeler et al. (1998) proposed a "police worldview" explanation of police deviance and highlighted the unique occupational subculture as well as the opportunity and organizational structural elements that facilitate the commission of deviant acts. Other examples of theoretically derived research include applications of opportunity theory (e.g., Alpert and Dunham, 1996),

the so-called "authoritarian personality" (Adorno, Frenkel-Brunswik, and Levinson, 1950; Neiderhoffer, 1967), the "rotten apple" explanation (Sherman, 1974), impulsivity (Pogarski and Piquero, 2003), and the police subculture perspective (Bayley and Mendelsohn, 1969; Stoddard, 1995; Van Maanen, 1978).

Research has also linked the ecological indicators of social disorganization and racial conflict to patterns of police misconduct, arguing that communities characterized by social network and cultural attenuation are those in which (1) police are most likely to take advantage of misconduct opportunities and (2) residents are least likely to have access to the conventional mechanisms of accountability (e.g., Kane, 2002). Additionally, several police scholars have tied the prevalence of police misconduct to organizational factors, such as variation in the vigor with which agencies pursue deviance (Sherman, 1978) and informal culture (Skolnick and Fyfe, 1993).[2]

Several researchers have conducted largely empirically driven examinations attempting to link static characteristics and life histories to police misconduct. For example, Waugh et al. (1988) reported that female officers in Queensland, Australia, were less likely than males to be subjects of complaints; this finding was supported by Hickman et al. (2000) in Philadelphia (but see Fyfe, Kane, Grasso, and Ansbro, 1998). Felkenes (1991) found that female officers' adherence to the "siege mentality" in the Los Angeles Police Department was substantively equivalent to male officers' adherence, which might suggest that organizational culture exerts similar effects across officer subgroups. Several studies have found that male officers are often more aggressive enforcers than are female officers (Bloch and Anderson, 1974; Forst, Lucianovic, and Cox, 1977; Melchionne, 1974; Sherman, 1975; Worden, 1989) and that male officers are more likely than females to use force in the course of their work (Grennan, 1987; Horvath, 1987).[3]

Although some research has found that African-American officers were more likely than White officers to engage in misconduct (e.g., Cohen and Chaiken, 1972, 1973), the link between race and misconduct has been confounded by both patterns of differential rule enforcement and assignments. In many urban police departments, African-American officers have been more commonly assigned to duties and geographies in which opportunities for misconduct are greater than average. Research on deadly force and police exposure to violence has also identified confounding variables that explain racial disparities in shooting rates (Bayley and Garafolo, 1989; Fyfe, 1980b; Geller and Karales, 1981). More recently, Fyfe et al. (1998) reported that Black officers' disparate rates of disciplinary action also were associated with differential patterns of assignment, rank, and off-duty behavior. Research has also examined the link between education and misconduct, showing mixed results (Eterno, 1996; Reuss-Ianni, 1983). Kappeler et al. (1992) reported that college-educated officers received fewer citizen complaints than non-college-educated officers. Truxillo, Bennett, and Collins (1998), however, found that—although

---

2. The current study is primarily interested in identifying individual-level factors associated with misconduct within the contexts of varying organizational and social ecological conditions. Therefore, we acknowledge research examining the community and organizational levels, but we focus our attention on research examining individual-level correlates.

3. It is unclear, however, that aggressive police posturing and the use of force actually represent misconduct.

college training was moderately associated with promotions in rank and supervisory ratings on job knowledge dimensions—neither 2-year nor 4-year college degrees were associated with disciplinary problems among officers.[4]

Some researchers have tested associations between life and career histories and occupational deviance in police. In one of the few large-scale studies of police misconduct, Cohen and Chaiken (1972, 1973) found that records of dismissal in prior jobs and military discipline were associated with internal police rules violations. Cohen and Chaiken's (1972, 1973) findings support a fundamental contention of control theorists (e.g., Gottfredson and Hirschi, 1990) that prior risky/deviant behavior predicts future risky/deviant behavior because of an individual's underlying low self-control. Cohen and Chaiken (1972, 1973) also found that officers who scored well in the police academy and on probationary evaluations were more likely than others to have advanced through the ranks and to win awards; and they were less likely to have engaged in misconduct or to have been disciplined (see also Grant and Grant, 1996). More recently, the Mollen Commission (1994) found a relatively high prevalence of prior arrests among suspended and dismissed officers, supporting Cohen and Chaiken's (1972, 1973) results.

Summary

Even in police departments characterized by systemic opportunities for corruption and abuse, it is likely that—although many officers might tolerate their colleagues' excesses and profit-motivated misconduct (Christopher Commission, 1991; Knapp Commission, 1972)—most officers do their work without engaging in (career-ending) misconduct. Thus, officers' individual characteristics, as well as their career experiences, might distinguish deviant officers from their colleagues. As noted, however, reliable empirical links between misconduct and individual-level factors are scarce. Indeed, although research has found some correlations between individual characteristics and coercive outcomes (i.e., use of force), research is more ambiguous when attempting to link individual factors to misconduct. The current study attempts to fill this gap. Although this study is not intended to be a rigorous test of a specific theory of crime or justice, its analyses and discussion of findings are informed by several etiological and organizational control perspectives.

Methods and Data

This study examined the personal and career histories of all 1,543 officers who were involuntarily separated from the NYPD for cause during 1975–1996. The first methodological step involved identifying the population of officers who met the *Bad Cops* study criteria. As a point of comparison, each study officer was then matched to a randomly selected academy classmate who had not been involuntarily separated. Note that the matching process did not stratify comparison officers on the basis of static (e.g., race or age) or other characteristics. The matches

4. The extent to which these findings are confounded by assignment factors is unclear, as evidence suggests that well-educated officers are likely to be moved out of street assignments and into staff, supervisory, and detective assignments where exposure to complaints is limited (Eterno, 1996; Reuss-Ianni, 1983).

were made strictly on the basis of academy class. Based on a review of prior research, we then developed research questions regarding the correlates of police misconduct and culled detailed information from the comprehensive NYPD personnel files on each study and comparison officer.[5] Bivariate and principal components analysis were first employed as data clarification and reduction techniques, and we then used multivariate binary and multinomial logistic regression to distinguish study officers from comparison officers along the specified set of independent variables while predicting the categorical dependent variable. Each of these methodological steps is described in more detail below.

Data Sources and Procedures

The first step in data collection involved identifying the population of officers separated from the NYPD between 1975 and 1996 for reasons of misconduct. At the time of the study, the NYPD maintained no central file or database that provided information on employment separations, whether for misconduct or other reasons. Instead, usually two to four times a week, the NYPD published personnel orders and disseminated them to every departmental unit. These orders reported every change in status (i.e., appointment, promotion, transfer, change in designation, resignation, retirement, vesting, dismissal, termination, or death) of sworn and non-sworn NYPD personnel. We gained access to all personnel orders, which were contained in multiple filing cabinets inside police headquarters. The personnel orders represented a starting point for the identification of separated officers.

In addition to reviewing these personnel orders, we accessed all other available NYPD record sources, most notably the personnel files held by the Personnel Records Unit (PRU). The most substantial part of the personnel file (known as the PA-15 in NYPD vernacular) was a lengthy application and information form completed by all police officer candidates and used as the basis for pre-employment investigations. Also from the PRU, we accessed files that contained officers' performance appraisals during recruit training, annual job performance evaluations, reports of disciplinary actions, information concerning changes in marital status, educational achievement, assignments, sick leave, injuries, and various other noteworthy events (e.g., commendatory letters, departmental recognition, serious vehicle accidents, and reports of line-of-duty injuries). Finally, we accessed and coded information from the Central Personnel Index (a computerized database that captures all critical career events), Management Information Systems Division, Internal Affairs Bureau, and the Department Advocate (the prosecutor in serious internal disciplinary actions). The only data to which the authors could not gain access involved the results of testing and interviews conducted by the NYPD's Psychological Services Unit—as these were protected by doctor–client confidentiality.

---

5.   Data collection was also guided by our review of prior research.

Defining Police Misconduct

Initially, the distinction between line-of-duty misconduct and that which was unrelated to officers' police status seemed reasonably well defined. After examining an initial sample of cases, however, we realized the difficulty of differentiating between police deviance and deviance committed by people who happen to be police officers. For example, many officers were dismissed because of their failure or refusal to take part in the NYPD's drug-testing program. Other officers were caught possessing or trafficking drugs. Such cases presented the insoluble problem of determining whether and to what extent these offenses were related to officers' membership in the NYPD. Unfortunately, examinations of NYPD files did not resolve these questions and, even had there been some practical way of locating and interviewing the officers involved, there is no reason to believe that they would be forthcoming about the details of their problems.

Given these concerns, we constructed our own classification scheme of police misconduct. Although it is informed by prior research, the classification scheme was expanded based on our review of the officers in this study and their misdeeds:

1. *Profit-motivated crimes*: All offenses, other than drug trafficking and whether on duty or off duty, in which the end or apparent goal of officers' wrongdoing doing was a profit

2. *Off-duty crimes against persons*: All assaultive behavior, except for profit-motivated robberies, by off-duty officers

3. *Off-duty public order crimes*: All offenses, other than drug trafficking or possession, against public order, including driving while intoxicated and disorderly conduct

4. *Drugs*: Possession and sale of drugs, and related conspiracies, as well as failing or refusing to submit to departmental drug tests

5. *On-duty abuse*: All offenses by on-duty officers that involved the use of excessive force, psychological abuse, or discrimination based on citizens' membership in a class (i.e., gender, race, ethnicity, or sexual preference)

6. *Obstruction of justice*: Conspiracy, perjury, official misconduct, and all offenses in which the apparent goal is obstruction or subversion of judicial proceedings

7. *Administrative/failure to perform*: Failure to abide by departmental regulations concerning attendance, performance, obedience, reporting, and other conduct not included in other offense types

8. *Conduct-related probationary failures*: All misconduct-related terminations of probationary officers in which misconduct in types 1–7 is not specified, and excluding simple failure in training programs

Identifying the Study and Comparison Officers, 1975–1996

A review of the personnel orders produced an original pool of approximately 3,000 officers who had career-ending events that might have been related to misconduct. We then reviewed the personnel history of each of those 3,000 officers to determine whether the cause for separation was related to misconduct.[6] We identified several categories of employment separation that met our criteria for inclusion in the study. First, "dismissal" is the term used to describe the firing of tenured civil servants in New York, including the police. Because dismissal is always a penalty for misconduct, all officers dismissed by the NYPD between 1975 and 1996 were included in the study. Second, the NYPD has traditionally followed a practice whereby it forced some officers to retire or resign under honorable conditions in return for their cooperation in investigating and prosecuting wrongdoing in which they had been a part. We reviewed the personnel orders and the personnel files of all officers whose resignation notices gave any indication that they had been in trouble when they left the department.[7] The lead author questioned internal affairs and trial room personnel for their recollections of any such cases.

In addition to identifying categories of inclusion, we also identified several categories of separation that excluded officers from the study. First, any officer who left the department in good standing (e.g., for employment with another police agency) was excluded from the study. Second, prior to the award of tenure, the NYPD typically does not dismiss officers because this category of involuntary separation requires some due process, beginning with the specification of charges. Instead, probationary officers typically are "terminated"; this designation requires only a statement from the Police Commissioner that an individual has shown to be an unsatisfactory probationer. Such terminations can occur because of misconduct or because of candidates' failures to satisfactorily meet the police academy's standards for performance in the academic, physical, or firearm and tactics training programs. If the termination occurred because of failures or inadequacies rather than misconduct, the officer was excluded from the analysis. Approximately 1,000 of the initial 3,000 cases were excluded because of performance failures. Probationary officers, however, whose terminations were obviously rooted in misconduct— including failure to abide by the police academy's disciplinary rules or, often, failure to take or pass a drug test—were included in the study.[8] The misconduct "screening" process produced a sample of 1,543 study officers.[9]

6.   For a detailed discussion of this procedure, the interested reader should see Fyfe and Kane (2006).

7.   One indicator was whether officers resigned "with the permission of the police commissioner," which is the NYPD's term of art for departures in good standing. We pulled and reviewed any file that indicated that officers had resigned without the commissioner's permission and included most of these cases in our files.

8.   In this process, we also encountered a small number of cases in which probationary officers had been decertified when it was discovered that they had concealed pre-employment histories of criminal behavior or mental illness. These cases, too, were excluded on the reasoning that, although these officers were effectively living a lie after they were hired, they would have been screened out in a more thorough pre-employment investigatory process.

9.   This data-collection strategy almost certainly missed a small number of forced resignations and retirements that were caused by misconduct. Nevertheless, we feel confident that we did everything reasonably possible to capture them all and that the few that we may have missed do not affect the direction or strength of our findings.

To generate a comparison group, we first identified the police academy classes to which all study officers belonged (based on start dates). Next, from police academy ledgers, we obtained rosters for all relevant academy classes (classes ranged from 1948 to 1996) and then matched each study officer to a randomly selected academy classmate.[10] This procedure allowed us to compare the study and comparison officers who had been hired under the same screening requirements, similar social and political contexts, and who likely had similar entry-level experiences with the organization.[11] In the end, we obtained a sample of 1,542 control officers.

Analytical Procedures

We engaged in a lengthy series of bivariate analyses (results shown in Table 2),[12] which served as a starting point for the multivariate analyses. Also—in an effort to protect against excessive multicollinearity, to reduce the number of variables in the models, and to identify discrete dimensions of officers' personal histories—independent variables associated with officers' criminal histories, social background conditions at appointment, and employment histories were subjected to a series of principal components analyses. Next, in an attempt to distinguish study officers from comparison officers along the specified set of predictor variables, we employed logistic regression modeling, which is useful for estimating the independent effects of a set of predictor variables on a binary outcome, while controlling for the influences of other covariates. The logistic regression analysis used all 3,085 cases and set the dependent variable to predict the study group (i.e., comparison officer = 0; study officer = 1).[13]

Finally, to examine the extent to which the multivariate relationships were stable across dimensions of misconduct, the study conducted a follow-up multinomial logistic regression analysis, specifying a four-category nominal-level dependent variable with the following coding scheme:

9. *Comparison officer* (*n* = 1,542): Not separated for misconduct
10. *Crime and serious misconduct* (*n* = 581): Police crime, profit-motivated serious misconduct, perjury, criminal abuse of authority, or other criminal events
11. *Administrative misconduct* (*n* = 430): Serious violations of departmental rules (e.g., associating with "known" criminal offenders, moonlighting as bouncers at nightclubs, making false statements in reports, and insubordination) or violations of more general rules and

---

10. For example, if there were 10 separated officers from the July 1980 recruit class, then 10 control officers were also selected randomly from that class.

11. There were several cases where a randomly selected comparison officer could not be used in the study. For example, a few comparison officers had resigned from the NYPD within the first few days after their appointments, so their files included insufficient information for comparisons of any kind. In other cases, the selected control officer had left the NYPD so long ago that his/her files had been destroyed in accord with the agency's 21-year document retention schedule. In all instances in which the original randomly selected control officer was determined to be unusable for analysis, the next available officer on the class roster was selected and included.

12. See Fyfe and Kane (2006).

13. All categorical variables were entered into the model using indicator contrasting, which designates a reference category against which the remaining categories are compared.

procedures (e.g., failure to maintain professional standards, failure to reside within the proper geographic distance of the city, failure to take action, or taking improper action)

12. *Drug test failures/refusals* (*n* = 420): Positive results on both random drug tests and drug tests that were "for cause" (e.g., for officers moving into special assignments or promotion or for officers whose behavior deviated sufficiently from "normal"). This category also includes officers who refused to take either random or for-cause drug tests (resulting in automatic dismissal)

The above categories are conceptually meaningful because they represent the primary dimensions of police misbehavior, but they are also complicated (and, hence, somewhat methodologically undermined) by a certain degree of overlap. For example, in some cases where an officer was arrested for drug dealing but not convicted, he was often dismissed for "associating with known drug dealers"—an administrative rule violation. In cases in which the official dismissal charge did not adequately represent the underlying behaviors that led to dismissal, we categorized the officer on the basis of the behaviors that led to the charge and not the charge itself.

The Research Setting: Misconduct and the NYPD, 1975–1996

As noted, the current study is roughly anchored by two major public scandals occurring approximately 20 years apart, each resulting in the formation of a Blue Ribbon Commission (i.e., the Knapp Commission in 1972 and the Mollen Commission in 1994) to study the causes and extent of the misconduct. The NYPD experienced several changes during the study period. Between 1975 and 1979, most police officers identified as corrupt during the Knapp Commission era were finally separated from service after often-lengthy investigations. It was also during this time that, because of citywide fiscal crisis, the NYPD laid off 3,000 officers and attrited by another 8,000—which shrank the department from 32,000 to 21,000 officers during a few month period (Kane, 2001). Between 1974 and 1979, no police officers were hired into the NYPD. From 1980 to 1984, when the city emerged from the fiscal crises, the NYPD hired new (and rehired previously employed) police officers quickly and with minimal screening to bring the department back to prefiscal crisis strength (Kane, 2001). Indeed, as police academy class ledgers indicate, between November 1979 and July 1984, the department hired 12,002 police officers.[14]

During the remainder of the study period, the NYPD continued to hire more police officers but with increasingly rigorous screening. In addition, the NYPD initiated a program in 1985 to screen all applicants and probationary police officers for drug use as part of the pre-employment selection process, during training at the police academy, and just prior to the expiration of the probationary period for the officers. Generally, a positive drug test results in dismissal from the department. In 1986, the NYPD expanded the Dole program and began testing employees

---

14. To a great extent, the rapid hiring of officers by the NYPD likely increased the variability of officers along key personal history indicators and static characteristics, countering the homogeneity problem that previous research has identified as problematic for the study of misconduct (McManus, 1969; Mollen Commission, 1994).

who were starting sensitive specialized assignments (e.g., Detective Bureau, Organized Crime Control Bureau, the Special Operations Division, and the Highway Unit). Finally, in 1989, the department introduced random drug testing of *all* in-service personnel assigned to the Organized Crime Control Bureau. By 1992, the NYPD doubled its annual random drug testing from 10% to 20% of all uniformed personnel. In 1993, the NYPD made drug testing a condition of promotion to sergeant (and later for lieutenant and captain; see Kane, 2001).

Our reasons for isolating this study period were both practical and theoretical. First, the identification and selection of study officers was an intensive, time-consuming process, and we weighed issues related to time and resources with need for a sufficient sample size.[15] Simply, selecting a starting point prior to 1975 would have been optimal, but practical concerns intervened. Practical considerations also influenced the selection of the end of the study. Specifically, our unique access to data allowed us to view sensitive information, and the NYPD sought to protect the anonymity of officers who were separated more recently as well as their supervisors who were most likely still working for the NYPD.

Although practical limitations were powerful influences, theoretical issues were also considered. It is likely that the changes that took place in the NYPD between 1975 and 1996 created an increasingly diverse police department (making the study of individual-level misconduct more methodologically reasonable), as well as altered the ways in which the organization responded to known deviance at different time periods. As Kane (2001) noted, during the fiscal crisis, many police officers alleged that it was difficult to commit an act of misconduct that resulted in separation because of the need of the agency to retain as many officers as possible. Moreover, from 1975 through the early 1990s, the department was relatively free from public scandals.[16] This is a key issue, for as long as an organization can remain relatively free from public scrutiny (i.e., between the anchor points of major scandals), some deviance could occur in the department without being sanctioned by the organization's external environment (Baker, 1996; Black, 1972). Under such conditions, the department is free to determine how to respond to the detected deviance.[17]

---

15. This project was funded by the National Institute of Justice (Award 96-IJ-CX-0053). As a result, we worked under the normal conditions associated with grant-funded projects, which included budget concerns, resource limitations, and deadlines.

16. Perhaps the largest occurred in 1986 when a group of officers from the 77th Precinct were implicated in a drug-related corruption scandal. Known as the "Buddy Boys," 38 officers were eventually disciplined for extorting money, as well as for using drugs both on and off duty (Kappeler et al., 1998). And, of course, the Mollen Commission issued its report in 1994.

17. Although this freedom operates within the boundaries of union, legal/collective bargaining, and administrative due process constraints.

Findings

The study's first major finding involves the prevalence of career-ending misconduct among of-ficers in the NYPD. During the study period, the NYPD averaged well over 30,000 uniformed officers per year, which ranged from 21,500 in the late 1970s to 38,000 at the end of the study period. On June 30,1975, the department employed more than 32,000 officers, and by the end of 1996, it had hired (or absorbed from the former Housing and Transit Authority police agencies) more than 45,000 additional officers. In all, the NYPD employed approximately 78,000 different individuals as police officers during the study period. Thus, the population of 1,543 officers separated for cause represents about 2% of all officers employed by the NYPD from 1975 to 1996. Figure 1 is a sequence plot showing the annual rates (per 1,000 officers) of misconduct-related separations from 1975 to 1996.

F I G U R E   1

Annual Misconduct-Related Separations in the NYPD
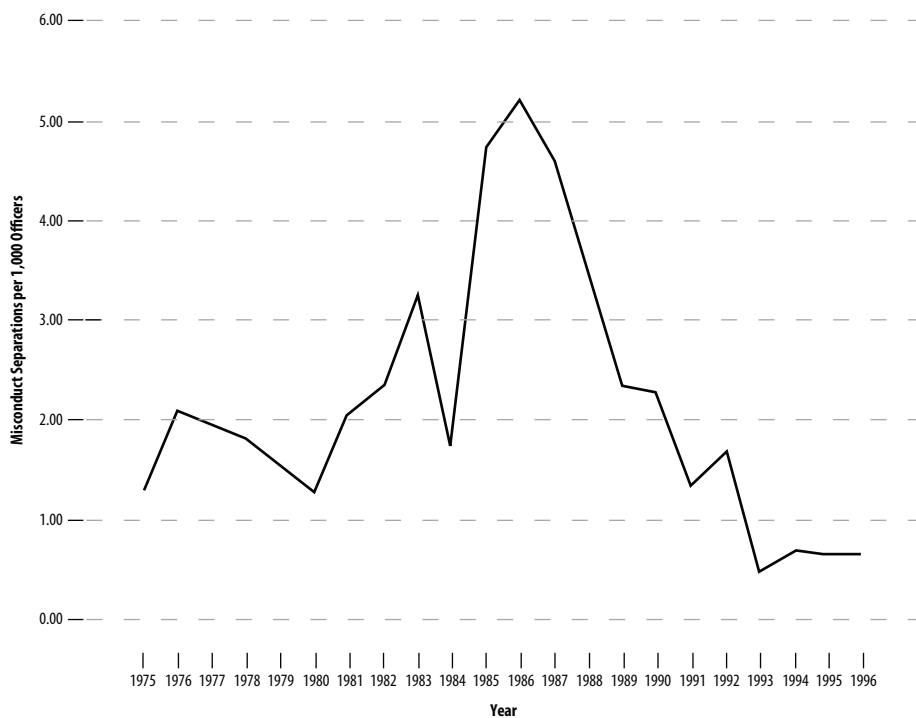
As the data in Figure 1 show, between 1976 and 1980, a slight but consistent decline oc-curred in the annual rates of career-ending misconduct in the NYPD. Recall that it was during this interval when the department instituted a hiring freeze and might have attempted to retain as many officers as possible. Once the department began hiring again (and it did quickly), an-

nual rates of misconduct began to rise, peaking in 1986 before starting to decline again. Recall that this was the year of the so-called "Buddy Boys" misconduct scandal (see footnote 16) when 38 officers were disciplined and/or separated all at once. In addition, 1986 was the year when the department implemented a fairly comprehensive drug-screening program, which may have led to a spike in drug-related misconduct separations. Perhaps most importantly, however, the NYPD increasingly hired officers with arrest records between 1983 and 1986. Indeed, this figure peaked in 1986 when 10% of all police recruits hired that year had arrest records. By 1989, 1.7% of all officers hired had arrest records. The Buddy Boys misconduct scandal, drug testing, and hiring officers with arrest records in combination might explain much of the spike in organizational career-ending misconduct that occurred during the mid-1980s. Interestingly, even as the department expanded its drug-testing program through the remainder of the 1980s and into the 1990s, annual rates of misconduct declined substantially.

Table 1 shows the charge specifications against the population of separated officers, employing our misconduct classification scheme. The 1,543 separated officers generated a total of 2,465 specific charges (approximately 1.6 charges per officer). Nearly one third of the charges involved administrative offenses, most typically for violation of departmental regulations concerning attendance, performance, obedience, and reporting. About one fifth of the charge specifications against officers involved drugs—either sale, possession, or failure of a departmental drug test (or refusal to submit a specimen), and 16% involved profit-motivated corruption, most commonly taking a bribe (25%), grand larceny (23%), and insurance fraud (12%). Other less common offenses include off-duty crimes (crimes against persons, 12%; public order offenses, 6%), obstruction of justice (11%), and conduct on probation (2%). Notably, only 119 of the separated officers ended their careers on charges that included on-duty abuse of authority (e.g., excessive force).

T A B L E   1

Charge Specifications against the Study Officers (*N* = 1,543), 1975–1996

Charge Specifications	Percent (*n*)
Administrative/failure to perform	30.1 (742)
Drugs	19.0 (468)
Profit-motivated crime	15.7 (387)
Off-duty crime against persons	11.6 (286)
Obstruction of justice	10.8 (266)
Off-duty public order crimes	5.8 (144)
On-duty abuse	4.8 (119)
Conduct on probation	2.2 (53)
Total	100.0 (2,465)

Characteristics of the Separated and Comparison Officers

Table 2 summarizes the life and career histories of the separated officers and their matched counterparts, and many significant differences are evident. With regard to personal history and background, study officers were more likely to be Black and Hispanic; were slightly more likely to be less than 22 years of age (at appointment); were more likely to be divorced, separated, or not married; were more likely to have children; were less likely to reside outside of the city; were more likely to have a prior criminal history; were less educated; were more likely to have negative prior employment experiences; and were more likely to have served in the military. With regard to their career history in the NYPD, study officers were more likely than their matched counterparts to have had the NYPD background investigator recommend against their hiring; more likely to have below standards performance ratings (and were less likely to have extremely or highly competent ratings from the police academy); more likely to have had command disciplines, formal department charges, and prior citizen complaints; more likely to have sick leaves and tardiness; and were less likely to have left patrol for promotion, investigative, or administrative units. In short, the bivariate analysis suggests that the study officers had numerous personal history, background, and employment characteristics that were presumably negative, and differentiated them from their colleagues who served without experiencing misconduct-related separation. The next section examines whether these characteristics withstand multivariate testing.

TABLE 2

Descriptive Comparison of Separated (Study) and Control Officers

Variable	Study Group (n = 1,543)	Comparison Group (n = 1,542)
Pre-Employment Personal History		
Female	15.1	13.4
White**	56.8	78.9
Black**	30.5	10.6
Hispanic*	11.9	9.6
Other	0.5	0.5
Born in New York City*	75.8	80.3
Born elsewhere in New York State**	5.7	9.0
Born in other U.S. state**	11.3	5.3
Born in Puerto Rico	1.9	2.3
Born outside the United States	3.2	2.5
Less than 22 years old at appointment*	27.7	24.4
Never married	64.0	66.7
Divorced/separated**	4.3	2.1
Children at appointment**	26.2	20.3
Resides outside New York City**	14.8	29.2
Arrested**	23.3	13.9
Violent crime arrests*	3.1	1.5
Property crime arrests**	7.6	3.6
Public order arrests**	14.6	5.0

JD/YO findings**	9.7	4.2
Misdemeanor convictions**	4.7	1.9
Moving violations*	60.8	55.2
Parking violations	28.9	30.8
Criminal/TAB summonses**	14.4	8.8
Own a motor vehicle**	46.2	53.9
Drivers license suspended**	17.9	11.8
GED or high school grad only**	17.2	12.5
Associate degree or higher**	6.4	14.7
Actively in school at appointment*	7.2	9.3
Unemployed for 30 days or more*	63.4	58.3
Professional or technical occupation*	5.7	8.3
Fired from job**	15.3	7.7
Negative comments by prior employer**	10.0	4.7
Disciplined by prior employer**	15.3	8.3
Workers' compensation claims	5.5	5.2
Previous civil service rejections**	15.3	9.3
Previously tested for NYPD	7.9	7.9
Prior police experience	3.6	4.4
Previously a transit officer pre-merge*	0.9	2.5
Previously a housing officer pre-merge	0.3	0.6
Merged from Transit Police in unification	1.0	1.0
Merged from Housing Police in unification	0.9	0.9
Previously an NYPD civilian employee	2.8	1.7
Previously an NYPD cadet/trainee	5.3	4.2
Owed more than $10,000 at appointment	3.3	4.1
Held home mortgage at appointment	1.9	2.6
Held vehicle loan at appointment	12.7	13.4
Held student loan at appointment	12.0	12.4
Held credit card debt at appointment	7.3	7.9
Held personal loan at appointment	15.1	13.4
Outstanding judgments at appointment*	1.7	0.7
Father an NYPD officer**	8.0	13.0
Members of immediate family arrested**	25.8	20.3
NYPD background investigator recommended not hiring**	15.4	7.0
Any NYPD staff recommended not hiring**	16.3	6.8
Resided in city when last known**	71.2	53.3
Married when last known**	42.0	65.5
Single when last known**	48.7	26.9
Children at separation**	31.0	23.7
Below-standards performance dimension rating**	21.7	12.4
Below-standards overall performance rating**	13.1	3.1
Military History		
Military service**	40.8	32.5
Among veterans only:		
Army veteran**	43.6	44.7
Navy veteran **	13.7	20.9
Marine veteran**	23.5	15.2
Air Force veteran*	11.9	16.7
Private/PFC/Corporal or seaman at discharge**	69.6	52.3

Sergeant/Petty officer at discharge**	29.5	40.5
Officer at discharge**	0.9	7.1
Court marshaled/disciplined**	21.4	14.4
Police Academy Performance		
Extremely/highly competent overall rating**	19.7	38.0
Academic average (mean)**	80.4	84.1
Extremely/highly competent law rating**	27.9	44.2
Extremely/highly competent police science rating**	22.0	38.0
Extremely/highly competent social science rating**	32.1	48.9
Physical education average (mean)**	86.6	89.9
Extremely/highly competent physical education rating**	42.3	56.1
Number of extremely/highly competent ratings**	2.07	3.22
Number of low/extremely low ratings**	0.50	0.20
Recruit demerits**	71.7	64.4
Command disciplines**	26.6	11.7
Formal department charges**	5.7	0.6
Non-line-of-duty sick leaves**	36.3	22.4
Line of duty sick leaves**	3.6	1.6
Latenesses**	26.1	18.9
Department History		
Prior complaints**	59.8	47.9
Assigned to patrol at separation**	59.3	51.1
Assigned to investigative unit at separation**	5.2	11.7
Assigned to admin staff job at separation**	7.1	13.1
Sergeant or higher at separation**	3.2	8.2
Notes. JD = juvenile delinquency.		

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

Principal Components Analysis

We used principal components analysis (PCA) as a data-reduction technique and as a protection against excessive multicollinearity in three categories of variables: criminal history, work history, and social conditions. Table 3 shows that the first PCA extracted three components from the model with all criminal history variables (explaining 47.4% of the cumulative model variance): a traditional criminal history dimension composed of violent, property, and misdemeanor crimes, as well as juvenile delinquency; a public order dimension composed of public order arrests and misdemeanor convictions; and a vehicle code violation dimension composed of traffic and parking summons and drivers' license suspensions.

Table 3 also shows the PCA for work history, from which two dimensions were extracted, explaining 49.1% of the variance. The first dimension reflects pre-employment disciplinary problems, including numbers of jobs from which an officer was fired, job-related disciplinary actions, and derogatory comments by prior employers. The second can be regarded as an employment reliability construct, composed of the number of workman compensation claims, jobs held, and 30-day (or longer) periods of unemployment. Last, two dimensions emerged

from the social conditions PCA (39.6% of variance explained): a conventional family construct composed of marital status and number of children and a "second start" dimension characterized by active enrollment in school and divorce or separation.

T A B L E   3

Principal Components Analysis of Criminal History Variables

Criminal History	EV = 2.81; % variance explained = 20.1	EV = 2.00; % variance explained = 14.3	EV = 1.82; % variance explained = 13.0
Violent crime arrests	.684	—	—
Property crime arrests	.703	—	—
Juvenile delinquency	.514	—	—
Misdemeanor convict	.496	.398	—
Public order arrests	—	.948	—
Moving violations	—	—	.732
Parking summonses	—	—	.627
Driver's license Suspension	—	—	.545

Work History	EV = 1.91; % variance explained = 31.9	EV = 1.04; % variance explained = 17.2	
Jobs fired from	.765	—	—
Number of disciplinary actions	.748	—	—
Derogatory comments from prior employer	.718	—	—
Workman comp claims	—	.525	—
Number jobs held	—	.640	—
30-day (or longer) period of unemployment	—	.685	—

Social Conditions	EV = 2.47; % variance explained = 27.5	EV = 1.09; % variance explained = 12.1	
Married	.916	—	—
Number of children	.704	—	—
Single (never married)	−.891	—	—
Active school enrollment	—	.546	—
Divorced	—	.706	—
Separated	—	.543	—
Living with partner	—	—	—

Notes. EV = Eigenvalue.

Logistic Regression Analysis

At step 0, the logistic procedure model produced a –2 log likelihood function of 2,128.61 and classified 51.6% of all cases correctly based on the intercept alone (i.e., essentially, no better than chance). At step 1 (during which all independent variables were simultaneously entered), the model produced a final –2 log likelihood function of 1,604.77 ($p < .001$) and correctly classified 86.5% of all cases. The final correct classification percentage coupled with the Cox and Snell *R*-squared and Nagelkerke *R*-squared (.55 & .73, respectively) suggests a highly internally reliable model with high predictive utility.

Table 4 shows that the model identifies risk and protective factors for career-ending misconduct.[18] While controlling for all other covariates, the strongest risk factors for misconduct were Black officer (*B* = 1.19), an average of one or more complaints per year of service (*B* = 1.11), and working in inspector precincts at time of incident leading to separation (*B* = 0.91). Specifically, Black officers were 3.27 times more likely than White officers to be dismissed for job-related misconduct; officers receiving an average of one or more complaints per year were 3.03 times more likely to be dismissed for misconduct compared with officers who received between zero and one complaint per year; and officers assigned to inspector precincts were 2.48 times more likely than officers assigned to staff units to be dismissed for occupational misconduct.[19] Other significant risk factors for police misconduct include being Latino (*B* = 0.69); being assigned to DI precincts[20] (*B* = 0.81), captain precincts (*B* = 0.71) or police academy and field training units (*B* = 0.61); and having criminal histories (*B* = 0.20), public order offense histories (*B* = 0.58), prior employment disciplinary problems (*B* = 0.32), and prior employee reliability problems (*B* = 0.15).

TABLE 4

Logistic Regression Estimates and Odds Ratios Predicting Police Misconduct

Variable	B	S.E.	Sig.	Odds
Dad NYPD Officer	−0.004	.001	.007	.996
Military Service [a]				
Army	−0.208	.175	.414	——
Navy	−0.235	.267	.786	——
Air Force	0.270	.262	.654	——
Marine Corps	0.087	.306	.432	——
Command Risk [a]				
Inspector precinct	0.909	.296	.002	2.48

18. The model correctly classified 86.5% of the cases, and produced a –2 log likelihood function of 1,604.77 (*p* < .001) with R-squared values of .55 and .73 (Cox and Snell and Nagelkerke, respectively).

19. This finding is likely a result of differential exposure and opportunity associated with patrol assignments. Officers working patrol in busy sectors have more opportunity to engage in career-ending misconduct than officers assigned to administrative and staff units.

20. Most NYPD precincts are commanded by captains, but several precincts are DI precincts, which are commanded by deputy inspectors. That both DI and captain precincts were significant suggest that patrol precincts in which officers work primarily "outside" are riskier than others.

DI precinct	0.813	.242	.001	2.26
Captain precinct	0.705	.225	.002	2.02
Police academy/FTU	0.605	.288	.036	1.83
Proactive investigation	−0.151	.476	.752	——
Detective/warrants	0.290	.316	.358	——
OCCB field units	0.830	.456	.069	——
Special patrol units	0.106	.299	.724	——
Supervisory rank	−0.639	.272	.019	.528
Age at appointment	−0.055	.002	.016	.946
Mean complaints per year	1.11	.250	.001	3.03
Educational Level at Appointment [a]				
College study (no degree)	−0.312	.143	.029	.732
Associate's degree	−1.241	.318	.001	.289
Bachelor's degree or higher	−0.799	.313	.011	.450
Other post-secondary	−0.218	.558	.696	——
Officer Sex	*−0.324*	*.203*	*.111*	——
Racial Composition [a]				
Black	1.185	.188	.001	3.27
Latino	0.692	.228	.002	2.00
Other minority	1.001	.755	.185	——
Prior Police Service	−0.241	.319	.355	——
Background Investigator Recommendation	0.259	.230	.259	——
Mayor at Time of Separation [a]				
Beame (1975–1977)	0.032	.220	.884	——
Koch (1978–1989)	−0.485	.288	.092	——
Dinkins (1990–1993)	−0.509	.281	.070	——
Criminal History				
Traditional dimension	0.199	.071	.005	1.22
Public order dimension	0.577	.083	.001	1.78
Vehicle code dimension	0.027	.065	.680	——
Years on the Job	−2.952	.261	.001	.052
Social Condition at Appointment				
Conventional family dimension	0.085	.198	.223	——
Second start dimension	0.140	.798	.068	——
Employment History				
Disciplinary problem dimension	0.318	.075	.001	1.37
Reliability dimension	0.149	.069	.031	1.16

Notes. FTU = Field Training Unit; OCCB = Organized Crime Control Bureau; S.E. = standard error.

[a]Reference categories are as follows: military service (none); command risk (staff unit); educational level (high-school diploma or less); racial composition (White); and mayor at separation (Giuliani).

As noted, the logistic regression model also found several factors that seemed to protect officers against police misconduct. The strongest of these were increased years on the job (B = –2.95), having an Associate's (B = –1.24) or a Bachelor's (B = –0.80) degree at time of appointment, and increased age at appointment (B = –0.06). In addition, officers whose fathers had served as NYPD officers (B = –0.004) were significantly less likely (albeit, not substantially)

than other officers to be dismissed for misconduct, and officers who achieved a supervisory rank ($B = -0.64$) were less likely than line officers to be dismissed for misconduct.

In addition to identifying risk and protective factors for career-ending police malpractice, the model also identified factors statistically unassociated with occupational deviance. Among these were military service, officer sex, prior police service, and background investigator recommendation. The model also included a variable that indicated the mayor at time of officer's separation to control for the effects of social and political climates in New York City, but it was a nonsignificant covariate of police misconduct.[21] Finally, assignment to a Proactive Investigation Unit (listed under Command risk) was not associated with career-ending misconduct.

Multinomial Logistic Regression Analysis

As noted, the study disaggregated the misconduct dependent variable into three substantive categories (crime/serious misconduct, administrative violations, and drug test failures/refusals) and one reference category (comparison officers), and it estimated a multinomial logistic regression model to observe any possible significant differences among types of career-ending misconduct. Model diagnostics showed that the correct classification percentages were 98.9 for the comparison officers, 79.0 for crime/serious misconduct, 50.0 for administrative violations, and 82.1 for drug test failures/refusals. The model correctly classified 83.3% of the cases and produced a composite pseudo R-squared of .82. These goodness-of-fit indicators show a high degree of internal reliability. Table 5 shows the results of the analysis. In the interest of brevity, only significant and other important variables are shown in Table 5.

TABLE 5

Multinomial Regression Estimates Predicting Three Categories of Police Misconduct

Variable	Crime/Serious Misconduct ($n = 581$)		Administrative Misconduct ($n = 430$)		Drug Test Failures or Refusals ($n = 420$)	
	B	S.E.	B	S.E.	B	S.E.
Dad NYPD Officer	−0.051**	.001	0.182	.668	0.186	.668
Command Risk[a]						
Inspector precinct	1.743**	.120	1.263**	.126	1.671**	.122
DI precinct	1.842**	.141	1.734**	.132	0.812*	.410
Captain precinct	1.890*	.167	1.530**	.142	1.433*	.672
Police academy/FTU	1.011*	.100	1.111*	.404	1.121	.711

21. In an earlier section, we briefly reviewed the social and political changes that occurred in both the NYPD and New York City throughout the study period. These changes were considerable. The lead author's prior work has explored the influence of some of these organizational and environmental changes on the prevalence of police misconduct (e.g., Kane, 2002). This article focuses on individual-level explanations of police misconduct, but we include the "mayor" variable as a control for the larger, contextual changes.

Proactive investigation	−0.251	.245	0.365	.350	0.261	.255
Detective/warrants	0.842	.613	0.985	.510	0.852	.838
OCCB field units	0.754	.533	0.657	.541	0.546	.412
Special patrol units	0.643	.379	0.512	.405	0.521	.409
Supervisory Rank	−0.633*	.303	0.532	.306	−0.549*	.171
Age at Appointment	−0.055*	.012	−0.112	.113	−1.470*	.254
Mean Complaints per Year	1.583**	.126	0.595*	.266	0.598*	.256
Educational Level at Appointment [a]						
College study (no degree)	−1.281*	.449	0.397	.475	0.316	.449
Associate's degree	−1.623**	.145	0.713	.477	0.654	.451
Bachelor's degree or higher	−1.724**	.125	−0.738	.548	−0.709**	.050
Other post-secondary	−0.876	.743	0.534	.345	0.521	.402
Racial Composition [a]						
Black	1.921**	.161	0.975	.751	0.878**	.124
Latino	1.640**	.091	0.717	.739	0.754**	.143
Other minority	1.001	.927	0.689	.620	0.623	.421
Criminal History						
Traditional dimension	1.072*	.234	1.061*	.483	1.071*	.484
Public order dimension	1.433*	.255	1.703	.972	1.622*	.554
Vehicle code dimension	0.027*	.007	1.684*	.612	0.276*	.127
Years on the Job	−0.790*	.400	−0.663	.413	−0.760*	.368
Employment History						
Disciplinary problem dimension	1.221**	.062	1.262*	.279	1.182*	.520
Reliability dimension	0.743*	.394	0.677*	.257	0.757	.731

Notes. FTU = Field Training Unit; OCCB = Organized Crime Control Bureau; S.E. = standard error.

[a] Reference categories are as follows: command risk (staff unit); educational level (high-school diploma or less); and racial composition (White). Reference category for outcome was comparison officers ($n = 1,542$).

*$p \leq .05$. **$p \leq .01$.

As the data in Table 5 show, the findings from the multinomial logistic model largely confirm those from the binary logistic regression model (shown in Table 4), with a few important exceptions. As with the binary logistic model, the multinomial model showed that previous arrests, previous employment problems (disciplinary troubles and unreliability), citizen complaints, working in a street-level assignment (i.e., a patrol precinct), and being a non-White officer were risk factors for criminal misconduct. Having a Bachelor's degree and making rank were protective factors. These findings largely held for administrative violations, except that race was not significantly (or substantively) related to administrative violations. In addition, education was not a protective factor for administrative violations. Interestingly, although it was not significant (but it holds for the present sample), officer rank was positively associated with administrative violations, whereas it was inversely associated with criminal or serious misconduct.

For drug failures, disciplinary problems at prior jobs (but not unreliability), citizen complaints, patrol assignments (but not drug enforcement assignments), and prior arrests were

significant, as well as race: Black and Hispanic officers were 2.41 and 2.12 times more likely, respectively, to be dismissed or otherwise separated for drug test failures/refusals as compared with White officers. Again, these findings largely confirm those for the criminal/serious misconduct outcome. In addition, a Bachelor's degree (but no other levels of education) was a significant protective factor against drug test failures and refusals.

Among all categories of misconduct problematic work histories, criminal justice contacts, and duty assignments where officers most likely interacted with the public on a regular basis were important risk factors. For the types of misconduct that were largely outside the department's control, that is, criminal misconduct (which generally gets referred to the District Attorney) and drug test failures (for which the department leaves itself virtually no discretion—drug test failures automatically lead to dismissal), Black and Latino officers were more likely to be separated for them. For administrative violations, where the department can largely choose how to investigate and respond, race was not a factor.

Discussion

This study examined the life and career histories of all officers who were dismissed or forced to leave the NYPD for reasons of misconduct during a 22-year period, and it compares them with a random sample of their police academy classmates who served without engaging in career-ending misconduct. Results suggest that career-ending misconduct is a relatively rare event in the NYPD, as the 1,543 study officers represent less than 2% of the 78,000 officers employed by the department during the study period. Moreover, the most common forms of misconduct involved administrative offenses and drugs, whereas separations for on-duty abuse were uncommon. Multivariate analyses demonstrated that many significant, identifiable differences were found between study and comparison officers that serve as risk and protective factors of police misconduct.

Limitations

Some limitations to the study warrant discussion. First, the study examines separated officers from one police department—the NYPD—which, given its size and unique history, presents problems for external validity. Second, the study relies on official data only. Despite our exhaustive efforts to identify and include all officers separated for reasons of misconduct, any misdeeds by NYPD officers that were not discovered or were not reported by the department are not addressed in this study. Similarly, this study focuses only on career-ending misconduct. Police malpractice that resulted in lesser forms of punishment (e.g., suspensions) were not included here.[22] Last, the study examines the time period from 1975 to 1996. As stated, the decision to limit our examination to this time period was based on both practical and theoretical considerations.

---

22.   The exception to this statement is that the study captures lesser forms of misconduct and discipline for those officers in the study—the separated officers and their matched counterparts.

The Nature and Prevalence Findings

Clearly, career-ending misconduct in the NYPD rarely occurs, which supports a body of literature indicating that most police officers in the United States perform their jobs without engaging in serious deviance. Among those who were separated from the department, perhaps the most surprising finding is the infrequency of separations for on-duty abuse: 119 officers, which represents 8% of separated officers and one tenth of 1% of officers employed during the study period (i.e., 119 of 78,000). For readers familiar with media accounts of the Louima and Diallo cases (which occurred outside the study period) and with allegations of widespread police abuse, this number might seem to be extremely small. Two explanations could apply. First, for several reasons, NYPD officers may use force less often than might be expected of urban police officers (see Chevigny, 1995; Domanick, 1994). Indeed, research suggests that the NYPD has traditionally used deadly force proportionally less often (e.g., per 1,000 officers) than is true of most other large police agencies (see, e.g., Geller and Scott, 1992; "Series on police use of force," 2001). Moreover, Skolnick and Fyfe (1993) reported that the NYPD has been the subject of civil rights complaints to the U.S. Department of Justice much less often (again, per 1,000 officers) than any of 58 other municipal and county agencies that had been the subject of a federal study on the issue (U.S. Department of Justice, 1992).

The second explanation involves the difficulty of proving allegations of police abuse. This involves the dilemma of demonstrating that officers who are entitled by law and policy to great discretion in use of force have crossed the barrier from reasonable to unreasonable force. This is a difficult burden of proof in criminal courts and police administrative proceedings (Cheh, 1996; Klockars, 1996). Consequently, findings of wrongdoing in either forum usually occur only when evidence is irrefutable—such as in the Rodney King and Abner Louima incidents. Furthermore, even if allegations of excessive force were sustained, it may be that during the period studied, the department tended to assign penalties that were less severe in nature than dismissal (e.g., 10-, 15-, or 30-day suspensions).

Risk Factors of Misconduct

The most intriguing finding from this study involves race and police misconduct. Although a multivariate analysis showed that Black officers were three times more likely to experience career-ending separation for serious misconduct and drug-test failures than were White officers, a closer inspection of the data highlights several other dimensions to the race–misconduct relationship. Early in the study period, non-White officers (Black, Hispanic, and Asian) were significantly more likely than white officers to be involuntarily separated for misconduct. Throughout the study period, this disparity seems to have faded where Hispanic and Asian officers are concerned such that the rates at which Hispanic and Asian officers were involuntarily separated became virtually indistinguishable from the separation rate for White officers. During the same period, the Black officers' rate decreased but remained much higher than those for other groups. Taken together, these trends suggest that, as the NYPD has become more diverse, it has become bet-

ter behaved.[23] This finding supports efforts to make police agencies closely representative of the populations they serve. The persistence of the finding for Black officers, however, remains troubling. This race finding has many possible, and not mutually exclusive, explanations. Martin (1980) studied the entry of women into the Washington, DC Metropolitan Police Department, and observed that they proceeded through a stage that Everett C. Hughes (1944) described as "tokenism." During this stage, members of new and growing groups are closely monitored by peers and supervisors in dominant groups, and they are treated with suspicion and a high degree of skepticism about their ability to perform. During this period, each token is treated as a representative of his or her group, and minor acts of misconduct may be viewed as symptoms of more serious problems. Under such close scrutiny, accounts of misconduct or substandard performance by individual tokens become "organizational lore" and are attributed to the entire group represented by the token. As token groups grow, however, dominants become more accustomed to, and less suspicious of, the presence and performance of tokens. With additional growth, the new group loses its token status, gains organizational power and prestige, and begins to exert real influence on the formal and informal cultures of the organization. At the same time, the solidarity of the dominant group breaks down as the new groups assimilate into it.

The data suggest that these processes have been at work in the NYPD. Hispanic and Asian officers, as well as women, have grown significantly in NYPD ranks during the last generation and have advanced rapidly through the ranks into positions of great influence. The representation of Hispanics as a percentage of the NYPD grew by 75.4% between 1986 and 1996. Similarly, Asians increased by 127.5% and women increased by 65.2%. African-American representation increased by only 28.8% (almost entirely as a result of the merger of the more heavily Black Housing and Transit Police Departments into the NYPD), whereas White representation decreased by 13.4%. In addition, the number of Hispanic supervisory and command personnel increased by 68.5% (from 257 to 433) between 1990 (the first year for which such data are available) and 1996, whereas Asian supervisors and commanders increased by 293.8% (16 to 63) and women supervisors and commanders increased by 111.4% (from 229 to 484). African-American supervisors increased by only 18.9% (from 333 to 396) during the same period.

By these measures, therefore, Hispanics, Asians, and women officers seem to have become well integrated into the NYPD during the study period. Their increased representation in the department might have allowed them to shed their status as tokens, and their movement up the ranks could have increased their influence over the department's culture and processes while also reducing their exposure to the risks of discipline faced by those at the department's lowest level. These processes of integration during the period studied may have left African-American officers as the department's "out group," with separation rates unlike those of the other racial groups or females.

---

23. This is true to the extent that "better" behaved is analogous to decreased rates of organizational misbehavior.

Although the current study cannot rule out the possibility of a race effect in individual cases, it seems unlikely that this disparity is the result of any quantifiable discrimination in the disposition of cases. Recall that the multinomial logistic model found that—although Black and Hispanic officers were more likely than Whites to be separated for serious/criminal misconduct and drug-test failures and refusals—they were no more likely than Whites to be separated for administrative misconduct. For criminal or other serious misconduct, as well as for drug test failures or refusing to take drug tests, the NYPD has little discretion over how to respond: As per departmental policy (and, in some cases, state law), such officers are generally dismissed from employment. The department has the *most* discretion over how to respond to administrative rule violations. If organizational responses to misconduct were racially biased, then we would expect those biases to show in the area of administrative dismissals/separations. As the multinomial findings show, however, race was not significantly (or substantively) associated with administrative misconduct.

Still, in this context as in others, race could be a proxy for some unmeasured factors, such as an urban experience or vigorous resistance to being charged. For example, Fyfe et al. (1998) found that although Black officers were more likely than White officers to be subjects of departmental discipline, they were also more likely than White officers to force a departmental trial in the adjudication of their disciplinary cases. Officers who forced departmental trials—regardless of racial background—were significantly more likely than officers who pled guilty (charges and specifications being equal) to be dismissed after a guilty case disposition. Thus, some percentage of Black officers' greater likelihood of "engaging" in career-ending misconduct in the current study could be an artifact of the way they often proceed through the disciplinary process.

Several other pre-employment factors emerged as significant predictors of misconduct, including prior criminal history and documented problems in prior jobs. Previous research has also documented these risk factors (Cohen and Chaiken, 1972, 1973; Mollen Commission, 1994). This finding is consistent with control theories (e.g., Gottfredson and Hirschi, 1990; Hirschi, 1969) suggesting that prior deviance predicts future deviance through low self-control. Thus, the results from this study suggest that departments' resource commitments to front-end screening is well founded. By screening out those with prior arrests and prior employment problems, departments can significantly reduce the likelihood of hiring future "bad cops."

Protective Factors of Police Misconduct

Among protective factors, the present findings highlight the importance of education and training. Evidence supporting college education requirements for police has been mixed, but the results here clearly show that officers with Associate or Bachelor's degrees were less likely to be separated for misconduct (i.e., criminal and drug failures) as compared with less educated officers. Moreover, training matters: Those who do well in the academy and during the probationary period are less likely to be separated than those who perform poorly at these early formative stages of their careers. These findings inform the long-standing debate about whether policing

is a profession in which traditional "classroom" training and education matters, or whether it is a craft in which the primary skills are learned through on-the-job training. The definitive answer is that both are vitally important. Further, the education and training findings—as well as the finding linking citizen complaints to misconduct—have implications for the use of early intervention or warning systems, as poor performance in these measures can serve as potential red flags (see also White, 2008). Walker (2005: 118) suggested that early intervention systems can "serve to identify and reward good performance," and education and training could serve as indicators in that capacity as well.

Factors Unrelated to Police Misconduct

The analysis also showed that many factors were unrelated to police misconduct, including officer sex, military service, and prior police service. Perhaps the most interesting of these nonsignificant findings involves assignment to Proactive Investigation. Many police scholars, dating back to August Vollmer, have argued that assignments in narcotics (historically) and street crimes units (more recently) are problematic because of the aggressive nature of such assignments, the deployment strategies of departments that use these assignments (especially street crimes units), and the degree to which officers (particularly in narcotics) must form close professional relationships with informants, drug dealers, and users under circumstances where supervision and accountability potential are low.

Many possible explanations exist for the nonsignificance of this variable. First, it could be that the misconduct of officers assigned to these units is more likely to go undetected or unreported because of the illegal activities of those who are victimized (e.g., drug dealers typically do not file complaints against officers who steal cash or drugs). Moreover, the misconduct might have been reported, but complaints were less likely to be sustained because of the credibility of the victims. Finally, the officers assigned to the aggressive, but elite, narcotics and street crimes units could have represented the best collection of officers serving in the NYPD during the study period. It could be that because of the extra layer of screening—the fact that candidates often must first establish track records of excellence in other assignments, and the competition to gain entrance to such coveted assignments—creates a sampling bias favoring the best-qualified officers. In fact, the current study cannot rule out any of these possible explanations.

Conclusions

The findings here confirm much of the conventional wisdom about police. Young officers who entered the police service with no post-secondary education, records of prior criminality, and prior poor employment; who did not advance in the police organization; who worked in busy patrol assignments; and who accumulated histories of complaints were more likely than others to have ended their careers in involuntary separation. Conversely, relatively well-educated officers with minimal or no criminal histories—perhaps including a family history in a police organization—and who advanced through departmental ranks were much less likely to engage in career-ending misconduct.

Perhaps the most salient policy implications of the present study relate to departmental screening processes. Because of the low visibility of police work, the unique opportunities for misconduct presented to police officers, and the conflict that often exists between the police and the public, it seems clear that police departments should continue to exclude people from policing who have demonstrated records of criminal involvement and employee disciplinary problems. These regulations represent evidence-based policy recommendations for which criminological perspectives developed for the general population (i.e., outside of policing) produced support (e.g., control theories, opportunity theories, and perhaps even routine activities theory). Moreover, the present findings also suggest the importance of *screening-in* (e.g., Grant & Grant, 1996) or otherwise identifying potential candidates whose presence in police departments could have the effect of making the organizations better behaved. Grant and Grant (1996) made this argument in terms of reducing the use of force in police departments, and to date, the current study is the only one identified that has produced empirical support for their arguments.

References

Adorno, Theodor W., Else Frenkel-Brunswik, and Daniel J. Levinson. 1950. *The authoritarian personality*. New York: Harper.

Alpert, Geoffrey and Roger Dunham. 1996. *Policing urban America*, 3rd ed. Prospect Heights, IL: Waveland Press.

Baker, Thomas. 1996. *Police ethics: Crisis in law enforcement*. Springfield: Charles C. Thomas.

Bayley, David H. and James Garofalo. 1989. The management of violence by police patrol officers. *Criminology*, 27: 1–25.

Bayley, David H. and Harold Mendelsohn. 1969. *Minorities and the police*. New York: MacMillan.

Black, Donald. 1972. The boundaries of legal sociology. *Yale Law Journal*, 81: 1086–1100.

Bloch, Peter B. and Deborah Anderson. 1974. *Policewomen on patrol: Final report*. Washington, DC: Police Foundation.

Bowker, Lee. 1980. A theory of the educational needs of law-enforcement officers. *Journal of Contemporary Criminal Justice*, 1: 17–24.

Carter, David. 1985. Hispanic perception of police performance: An empirical assessment. *Journal of Criminal Justice*, 13: 487–500.

Chappell, Allison T. and Alex R. Piquero. 2004. Applying social learning theory to police misconduct. *Deviant Behavior*, 25: 89–108.

Cheh, Mary. 1996. Are lawsuits the answer to police brutality? In (William Geller and Hans Toch, eds.), *Police violence: Understanding and controlling police abuse of force*. New Haven, CT: Yale University Press.

Chevigny, Paul. 1969. *Police power: Police abuses in New York City*. New York: Pantheon Books.

Chevigny, Paul. 1995. *Edge of the knife: Police violence in the Americas*. New York: Free Press.

Christopher Commission. 1991. *Report of the Independent Commission on the Los Angeles Police Department.* Los Angeles, CA: Christopher Commission.

Cohen, Bernard and Jan M. Chaiken. 1972. *Police background characteristics and performance.* New York: RAND Institute.

Cohen, Bernard and Jan M. Chaiken. 1973. *Police background characteristics and performance.* Lexington, MA: Lexington Books.

Domanick, Joe. 1994. *To protect and serve: The LAPD's century of war in the city of dreams.* New York: Pocket Books.

Eterno, John. 1996. *Cadets and policing: An analysis of the New York City Police Department's Cadet Corps.* Academy of Criminal Justice Sciences Annual Meeting, Las Vegas, NV.

Felkenes, George. 1991. Affirmative action in the Los Angeles Police Department. *Journal of Police Science and Administration*, 15: 138–152.

Forst, Brian, Judith Lucianovic, and Sara J. Cox. 1977. *What happens after arrest? A court perspective of police operations in the District of Columbia.* Washington, DC: Institute for Law and Social Research.

Fyfe, James J. 1980a. Always prepared: Police off-duty guns. *Annals of the American Academy of Political and Social Science*, 452: 72–81.

Fyfe, James J. 1980b. *Philadelphia police shootings, 1975–78: A system model analysis.* Washington, DC: Report for the Civil Rights Division, U.S. Department of Justice.

Fyfe, James J. 1986. The split-second syndrome and other determinants of police violence. In (Anne T. Campbell and John J. Gibbs, eds.), *Violent transactions.* Oxford, UK: Basil Blackwell.

Fyfe, James J. 1987. Police shooting: Environment and license. In (Joseph E. Scott and Travis Hirschi, eds.), *Controversial issues in crime and justice.* Newbury Park, CA: Sage.

Fyfe, James J. 1988. Police use of deadly force: Research and reform. *Justice Quarterly*, 5: 165–205.

Fyfe, James J. and Robert J. Kane. 2006. *Bad cops: A study of career-ending misconduct among New York City police officers.* Final Report. Washington, DC: National Institute of Justice.

Fyfe, James J., Robert J. Kane, George Grasso, and Michael Ansbro. 1998. *Race, gender, and discipline in the New York City Police Department.* Paper presented at the Annual Meeting of the American Society of Criminology, Washington, DC.

Geller, William A. and Kevin Karales. 1981. Shootings of and by Chicago police: Uncommon crises—Part I shootings by Chicago police. *Journal of Criminal Law and Criminology*, 72: 1813–1866.

Geller, William A. and Michael S. Scott. 1992. *Deadly force: What we know.* Washington, DC: Police Executive Research Forum.

Gottfredson, Michael R. and Travis Hirschi. 1990. *A general theory of crime.* Stanford, CA: Stanford University Press.

Grant, J. Douglas and Joan Grant. 1996. Officer selection and the prevention of abuse of force. In (William Geller, and Hans Toch, eds.), *Police violence: Understanding and controlling police abuse of force.* New Haven: Yale University Press.

Grennan, Sean A. 1987. Findings on the role of officer gender in violent encounters with citizens. *Journal of Police Science and Administration*, 15: 78–85.

Hale, Donna. 1989. Ideology of police misbehavior: Analysis and recommendations. *Journal of Ideology*, 13: 59–85.

Hickman, Matthew J., Alex R. Piquero, and Jack R. Greene. 2000. Discretion and gender disproportionality in police disciplinary systems. *Policing: An International Journal of Police Strategies & Management*, 23: 105–116.

Hirschi. Travis. 1969. *Causes of delinquency*. Berkeley: University of California Press.

Horvath, Frank. 1987. The police use of deadly force: A description of selected characteristics of intra-state incidents. *Journal of Police Science and Administration*, 15: 226–228.

Hughes, Everett C. 1944. Dilemmas and contradictions of status. *American Journal of Sociology*, 50: 353–359.

Kane, Robert J. 2001. *Urban ecology and police malpractice: Identifying contexts for career-ending misconduct in the New York City Police Department*. Unpublished dissertation, Temple University.

Kane, Robert J. 2002. The social ecology of police misconduct. *Criminology,* 40: 867–896.

Kappeler, Victor E., Alan D. Sapp, and David L. Carter. 1992. Police officer higher education, citizen complaints and departmental rule violations. *American Journal of Police*, 11: 37–54.

Kappeler, Victor, Richard D. Sluder, and Geoffrey P. Alpert. 1998. *Forces of deviance: Understanding the dark side of policing*. Prospect Heights, IL: Waveland Press.

Klockars, Carl. 1996. A theory of excessive use of force and its control. In (William Geller and Hans Toch, eds.), *Police violence: Understanding and controlling police abuse of force.* New Haven: Yale University Press.

Knapp Commission. 1972. *Report of the New York City Commission to investigate allegations of police corruption and the city's anti-corruption procedures*. New York: Bar Press.

Martin, Susan E. 1980. *Breaking and entering: Policewomen on patrol*. Berkeley: University of California Press.

McManus, George E. 1969. *Police training and performance study*. Washington, DC: National Institute of Law Enforcement and Criminal Justice.

McMullan, M. 1961. A theory of corruption. *Sociological Review*, 9: 181–201.

Melchionne, Teresa M. 1974. The changing role of policewomen. *Police Journal*, 47: 340–358.

Mollen Commission. 1994. *Anatomy of failure, a path for success: The report of the Commission to investigate allegations of police corruption and the anti-corruption procedures of the New York City Police Department*. New York: City of New York.

Muir, William K. 1977. *Police: Street corner politicians*. Chicago, IL: University of Chicago Press.

Neiderhoffer, Arthur. 1967. *Behind the shield*. Garden City, NY: Doubleday.

Pogarsky, Greg and Alex R. Piquero. 2003. Studying the reach of deterrence: Can deterrence theory help explain police misconduct? *Journal of Criminal Justice*, 32: 371–386.

Reuss-Ianni, Elizabeth. 1983. *Two cultures of policing: Street cops and management cops*. New Brunswick, NJ: Transaction.

"Series on police use of force in Prince George's County, MD." 2001. *Washington Post*, July 1–4. Retrieved on August 1, 2001 from washingtonpost.com/wp-srv/content/nation/investigative/policeshootings.html.

Sherman, Lawrence. 1974. *Police corruption: A sociological perspective*. Garden City, NY: Anchor Books.

Sherman, Lawrence. 1978. *Scandal and reform: Controlling police corruption*. Berkeley: University of California Press.

Sherman, Lewis J. 1975. An evaluation of policewomen on patrol in a suburban police department. *Journal of Police Science and Administration*, 3: 434–438.

Skolnick, Jerome H. and James J. Fyfe. 1993. *Above the law: Police and the excessive use of force*. New York: Free Press.

Stoddard, Ellwyn R. 1995. The informal "code" of police deviancy: A group approach to "blue-coat crime." In (Victor Kappeler, ed.), *Police and society: Touchstone readings*. Prospect Heights, IL: Waveland Press.

Toch, Hans. 1996. The violence-prone officer. In (William Geller and Hans Toch, eds.), *Police violence: Understanding and controlling police abuse of force.* New Haven: Yale University Press.

Truxillo, Donald M., Suzanne R. Bennett, and Michelle L. Collins. 1998. College education and police job performance: A ten-year study. *Public Personnel Management*, 27: 269–280.

U.S. Department of Justice. 1992. *Police brutality study: FY1985–FY1990*. Washington, DC: Author.

Van Maanen, John. 1978. The asshole. In (Peter K. Manning and John Van Maanen, eds.), *Policing: A view from the street*. Santa Monica, CA: Goodyear Publishing.

Walker, Samuel. 2005. *The new world of police accountability*. Thousand Oaks, CA: Sage.

Waugh, Linda, Andrew Ede, and Avril Alley. 1988. Police culture, women police, and attitudes towards misconduct. *Policing: An International Journal of Police Science & Management*, 1: 288–330.

White, Michael D. 2008. Identifying good cops early: Predicting recruit performance in the academy. *Police Quarterly*, 11: 27–49.

Worden, Robert E. 1989. Situational and attitudinal explanations of police behavior: A theoretical reappraisal and empirical assessment. *Law & Society Review*, 23: 667–711.

Robert J. Kane, Ph.D., is an associate professor in the School of Criminology and Criminal Justice at Arizona State University. Kane's primary areas of research are police authority and accountability, the urban ecology of policing, and the social ecology of crime and violence in urban communities. Kane's work has been published in *Criminology, Justice Quarterly, Criminology & Public Policy, Crime & Delinquency, Criminal Justice and Behavior*, among other outlets.

Michael D. White, Ph.D., is an associate professor in the School of Criminology and Criminal Justice at Arizona State University. White's primary research interests involve the police, including use of force, misconduct, training, and performance measurement. White's recent work has been published *in Criminal Justice and Behavior, Journal of Experimental Criminology*, and *Crime & Delinquency*.

Police officer misconduct as normal accidents
An organizational perspective

William R. King

S a m   H o u s t o n   S t a t e   U n i v e r s i t y

Much has been written about police officer misconduct (cf. Kappeler, Sluder, and Alpert, 1994; Klockars, Ivkovich, and Haberfeld, 2004), and I think it is fruitful to consider misconduct as the result of three factors that are as follows: individual officer attributes, community or ecological factors, and organizational factors. Kane and White (2009, this issue) and Kane (2002) have addressed the first two issues, so I will comment on two organizational factors, which could be manipulated to curb police officer misconduct. My discussion integrates structural contingency theory, through administrative rule making, with the literature on normal accidents and high-reliability organizations (HROs). First, I will present a quick review of the three dimensions of police officer misconduct.

Attributes of Police Officers: Bad Apples

Kane and White's (2009) groundbreaking study of career-ending police officer misconduct in the New York City Police Department (NYPD) is presented in the context of individual officer attributes and assignments. Simply put, bad cops share a handful of common attributes that are evident before their employment as police officers. This individual-level perspective fits well with Sanders' (2008) research on the personal attributes of "good cops" and Sherman's (1974) notion of bad apples.

Environment and Community Factors: Bad Orchards

Environment- and community-level variables also are related to officer misconduct. The amount and types of trouble police officers can become involved in is partially related to their assignment and the area in which they work (Sherman, 1978). If there are no drug dealers, then an officer cannot rob drug dealers. At a higher, or macro, level of analysis, Kane's (2002) research with officer misconduct in the NYPD indicates that certain community characteristics, or changes

in these characteristics (percent Latino population, population mobility, and disadvantage), are related to misconduct. Kane's (2002) findings substantiate police officer behavior research, which found that community characteristics affected police officer arrest decisions (Brown and Frank, 2006; Smith, Visher, and Davidson, 1984).

Organizational Correlates: Bad Barrels

What is missing from this picture of police corruption is a third group of factors related to police officer misconduct: organizational attributes. Little empirical research exists that concerns the organizational correlates of police misconduct, but scholars have noted the importance of police organizations in preventing or promoting misconduct (Klockars et al., 2004; Sherman, 1978). In particular, Klockars et al. (2004) posited that police organizations can prevent misconduct by (1) creating rules, (2) creating surveillance mechanisms to detect misconduct and then punish wrongdoers, and (3) dismantling the blue wall of silence. The first two preventive measures (rule making and surveillance) concur with the literature concerning the control of police officer discretion (Walker, 1993), and I have suggested using these reforms to reduce police dumping (King and Dunn, 2004). Solid evidence exists that, when properly implemented, rule making can control substantial police behaviors, such as the decision to arrest and to use deadly force. So, good reason exists to expect that some organizational attributes affect overall levels of police officer misconduct.

These two preventive policies (rule making and surveillance) are structural, formal, and hard attributes, which root them firmly in structural contingency theory (Langworthy, 1986; Maguire, 2003). The policies are structural because they entail changing the formal organizational structure, such as creating internal organizational units, which will surveil officers, and early intervention or early warning systems. These changes are formal, because they add rules or use scheduling arrangements to weaken the informal relationships between officers and between officers and their communities. Finally, the changes are hard because they are visible to organizational members and inquisitive outsiders. For instance, special units and formal rules will be visible on the organizational chart, in the written rules, and in annual reports.[1]

Organizational rule making and surveillance, provided they are properly implemented, show great promise in curbing objectionable officer behaviors. The next few decades will likely see this innovation proliferate throughout the policing industry. In some instances, rule making and surveillance will reap great rewards, such as when the unacceptable behavior or its outcomes are visible (e.g., use of deadly force), the rules are created at the agency level, the officers buy into the benefit of the rule, the incidents of rule breaking are reviewed, and the wrongdoers are held accountable (King and Dunn, 2004; Walker, 1993).

---

1. Skeptics of police reform, and some organizational and institutional theorists, will mutter that these structural contingency reforms are really just empty charades enacted by police agencies to signal powerful sovereigns that the agency is responding appropriately to officer misconduct.

Future Organizational Solutions to Officer Misconduct:

High-Reliability Organizations

At some point, the benefits of rules will plateau. Some objectionable officer behaviors will persist, even if at low rates. Additional rules or strengthening the elements of successful rule making will not quash the problem. Officer misconduct will persist, and civil suits against officers and agencies will continue to be won by plaintiffs wrongfully harmed by the police. For some people, these problems will be rare and will be deemed acceptable and unavoidable costs of doing business. For others, such events will remain unacceptable.

At this future point, policymakers should turn their attention to organizations that are good at avoiding events called *accidents*, which are—by anyone's accounting—tragic, harmful, dreadful events.[2] Examples of accidents include serious mishaps at nuclear power plants (Perrow, 1999), the deaths of fire and emergency medical services personnel at critical incidents (Weick, 1993), cases of fratricide or friendly fire in the military (Naylor, 2005; Snook, 2000), and disasters involving space shuttles (Starbuck and Farjoun, 2005).

HROs, a class of organizations, work in dynamic, turbulent environments in which mistakes can compound, cascade quickly, and produce accidents (Weick and Sutcliffe, 2001; Weick, Sutcliffe, and Obstfeld, 1999). Despite working under such demanding conditions, HROs are remarkably good at preventing organizational accidents, and they resort only partially to rule making and surveillance to avoid accidents. What elements of HROs might we use to prevent police misconduct if we conceptualize police officer misconduct as an organizational accident?

The HRO literature identifies one element of organizations as especially germane to policing that can prevent accidents: organizational culture. HROs rely heavily on a special organizational culture to prevent accidents. This culture is focused on mindfulness, which is defined as a "preoccupation with failure, reluctance to simplify interpretations, sensitivity to operations, commitment to resilience, and deference to expertise" (Weick and Sutcliffe, 2001: 10). To put it simply, HROs ensure that all employees are focused on detecting small aberrations in the environment, performance, or conditions, and employees are empowered to act when they encounter any of these aberrations. Small discrepancies can quickly grow into serious problems and accidents if left unattended, and line-level employees are often in the best position to identify the problem and address it before it grows into an accident.

Compared with rule making and surveillance, the cultural solution of mindfulness is soft, informal, and cultural. Mindfulness is soft and informal because it does not rely on formal, written rules, nor specialization of organizational structure to be effective. Mindfulness does not appear on an organizational chart nor in a rule book, and it is not easily evident to an external observer. Instead, mindfulness works at a cultural level, by changing the "attitudes and values

---

2.   Many types of police officer misconduct involve the willing involvement of officers; "they want to do it," which in turn suggests culpability and invokes a criminological perspective. Normal accidents have no such willing participants because they are tragic accidents in the true sense of the word. For the reader to follow this stream of thought, it is best to suspend their concern with officer culpability in misconduct.

that are shared and socially transmitted among groups of people" (Paoline, 2001: 7). I envision that mindfulness in a police agency would empower officers to act quickly when they see that a fellow officer is headed for trouble. If pre-employment indicators provide forewarning that some officers will engage in misconduct, then it is likely that problem officers continue to exhibit other troublesome indicators at work before they engage in more serious acts of misconduct.

My perspective on misconduct is influenced partially by my experiences working as a uniformed and plain-clothes security supervisor for a police agency in Ohio. It was common knowledge which officers were goofing off and creating problems. A small minority of these officers would eventually end up in more serious trouble, and some would be released from employment. Looking back at the work history of those individuals who were released revealed no surprises. Multiple indicators were present that were evident to anyone who worked alongside these troublesome officers.

This knowledge of troublesome behaviors is akin to knowledge of small aberrations in performance or environment in the HRO literature. Line-level officers are in the best position to preemptively identify officers destined for career-ending misconduct. If an HRO approach to preventing misconduct is to be successful, then two issues must be addressed. First, my contention that troublesome officers are known to most line officers before they make a serious mistake must be verified empirically. Second, line officers must be empowered to correct, coach, and reign in bad officers. Using line officers as coaches aligns with the notion of policing as a craft and should appeal to officers who take pride in being skilled operators. But, it is a proposal fraught with potential problems (even accidents of its own), which must be worked out by mindfulness on the parts of those implementing the solution.

Applying Lessons from the NYPD

Applying lessons from the NYPD to another agency is a bit like wearing someone else's shoes; sometimes they fit well, and sometimes they do not. Two useful lessons exist that one should bear in mind about the present study and what police executives can learn from the NYPD.

The first lesson is a general caution about learning from the NYPD. The NYPD is unlike the majority of police agencies in the United States because of its size; it is an organizational behemoth. With 53,029 full-time employees (in 2000; U.S. Department of Justice, 2000), the NYPD is more than three times the size of the next largest U.S. police agency, the Chicago Police Department.[3]

Furthermore, more than half of all U.S. police agencies have 12 or fewer full-time employees (U.S. Department of Justice, 2000). Police executives from other nations, especially nations with large national or regional police agencies, can more readily apply lessons from the NYPD to their agencies. But most police chiefs in the United States should be wary of adopting NYPD innovations or policies without serious forethought.

---

3. If the 9,253 smallest U.S. police agencies were merged into 1 police agency (these are all agencies with between 1 and 15 full-time employees), this new agency (with 50,278 full-time employees) would still not eclipse the size of the NYPD.

Despite this warning, some findings from the NYPD can be applied to other police agencies. Kane and White's (2009) research is one such study and few attributes exist in this study that are unique to New York—including the research methodology, the officers, and the findings. The findings provide useful information for police executives in most police agencies, no matter their size.

The second lesson concerns the willingness of the NYPD to open its records to researchers in order to facilitate research on an important topic. The NYPD should be lauded for providing access for Kane and White's (2009) project. I hope other police executives will follow suit, for good record keeping coupled with access to researchers yields great benefits for the policing industry.

References

Brown, Robert A. and Frank, James. 2006. Race and officer decision making: Examining differences in arrest outcomes between Black and White officers. *Justice Quarterly*, 23: 196–126.

Kane, Robert J. 2002. The social ecology of police misconduct. *Criminology*, 40: 867–896.

Kane, Robert J. and Michael D. White. 2009. Bad cops: A study of career-ending misconduct among New York City police officers. *Criminology & Public Policy*. This issue.

Kappeler, Victor E., Richard D. Sluder, and Geoffrey P. Alpert. 1994. *Forces of deviance: The dark side of policing*. Prospect Heights, IL: Waveland Press.

King, William R. and Thomas Dunn. 2004. Police initiated trans-jurisdictional transport of troublesome persons. *Police Quarterly*, 7: 339–358.

Klockars, Carl B., Sanja Kutnjak Ivkovich, and Mary R. Haberfeld. 2004. *The contours of police integrity*. Thousand Oaks, CA: Sage.

Langworthy, Robert H. 1986. *The structure of police organizations*. Westport, CT: Praeger.

Maguire, Edward R. 2003. *Organizational structure in American police agencies: Context, complexity, and control*. New York: SUNY Press.

Naylor, Sean. 2005. *Not a good day to die: The untold story of Operation Anaconda*. New York: Berkley Caliber.

Paoline, Eugene A., III. 2001. *Rethinking police culture: Officers' occupational attitudes*. New York: LFB Scholarly Publishing.

Perrow, Charles. 1999. *Normal accidents: Living with high-risk technologies*. Princeton, NJ: Princeton University Press.

Sanders, Beth A. 2008. Using personality traits to predict police officer performance. *Policing: An International Journal of Police Strategies & Management*, 31: 129–147.

Sherman, Lawrence W. (ed.). 1974. *Controlling police corruption: A sociological perspective*. Garden City, NY: Anchor.

Sherman, Lawrence W. 1978. *Scandal and reform: Controlling police corrupton*. Los Angeles, CA: University of California Press.

Smith, Douglas A., Christy A. Visher, and Laura A. Davidson. 1984. Equity and discretionary justice: The influence of race on police arrest decisions. *The Journal of Criminal Law & Criminology*, 75: 234–249.

Snook, Scott A. 2000. *Friendly fire: The accidental shootdown of U.S. blackhawks over northern Iraq*. Princeton, NJ: Princeton University Press.

Starbuck, William H. and Moshe Farjoun (eds.). 2005. *Organization at the limit: Lessons from the Columbia disaster*. Malden, MA: Blackwell Scientific.

U.S. Department of Justice. 2002. *Census of state and local law-enforcement agencies (CSLLEA), 2000* [United States: Computer file]. Conducted by U.S. Department of Commerce, Bureau of the Census. ICPSR, ed. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [producer and distributor].

Walker, Samuel. 1993. *Taming the system: The control of discretion in criminal justice, 1950–1990*. New York: Oxford University Press.

Weick, Karl E. 1993. The collapse of sensemaking in organizations: The Mann Gulch disaster. *Administrative Science Quarterly*, 38: 628–652.

Weick, Karl E. and Kathleen M. Sutcliffe. 2001. *Managing the unexpected: Assuring high performance in an age of complexity*. San Francisco, CA: Jossey Bass.

Weick, Karl E., Kathleen M. Sutcliffe, and David Obstfeld. 1999. Organizing for high reliability: Processes of collective mindfulness. In (Barry M. Staw and Robert I. Sutton, eds.), *Research in Organizational Behavior*. Greenwich, CT: JAI.

William R. King, Ph.D., received his degree from the University of Cincinnati in 1998 and is an associate professor of criminal justice at Sam Houston State University. His primary research interests involve change and continuity in police organizations. This research focus encompasses studying trends in police–civilian employees, long-term structural change in U.S. police organizations, police hierarchy and rank structure, and the adoption of innovation. King also is interested in assessing the effectiveness of systems that process forensic evidence and homicide investigation. From 2005 to 2008, he was involved in assessing the organizational capacity of agencies in the nation of Trinidad and Tobago to process physical evidence and to respond to violent crime.

P O L I C E   M I S C O N D U C T

Rotten apples, rotten branches, and rotten orchards

A cautionary tale of police misconduct

Sanja Kutnjak Ivković

Michigan State University

Police misconduct is a complex, multilayered phenomenon. As many cases have demonstrated, some police officers abuse the trust that society has bestowed upon them and become notorious law violators: They lie, steal, accept bribes, rob drug dealers, sell drugs, and turn a blind eye when they see other police officers beating up a suspect (e.g., Mollen Commission, 1994). The consequences of such acts reach far beyond the individual police officer and potentially result in the loss of legitimacy of, and confidence in, the police organization at large.

The Kane and White article (2009, this issue) taps into this complex world by exploring the correlates of "career-ending misconduct"—police misconduct detected by the authorities that ultimately leads to the application of the most severe disciplinary option—namely, the dismissal of the police officer. Kane and White report findings consistent with what scholars (see Greisinger, Slovak, and Molkup, 1979; Malouff and Schutte, 1986) and police administrators have intuitively suspected for some time and what a multitude of actual cases support (Burns and Sechrest, 1992; Delattre, 1989; Human Rights Watch, 1998; Los Angeles Police Department [LAPD], 2000): Hiring applicants under relaxed hiring standards could backfire. Specifically, Kane and White report that police officers who had criminal history and employment disciplinary problems at the beginning of their service with the New York Police Department (NYPD) were more likely to be fired for serious misconduct than the police officers without criminal history or employment disciplinary problems. Similarly, Kane and White report a negative relation between applicant education and the likelihood of subsequently being fired for rule-breaking behavior.

Direct correspondence to Sanja Kutnjak Ivković, School of Criminial Justice, Michigan State University, 560 Baker Hall, East Lansing, MI 48824 (e-mail: kutnjak@msu.edu).

Policy implications of this research seem clear: Efforts should be concentrated at the front end of the police officers' careers with thorough, intense, and complete background checks of applicants. Standards such as no criminal record, no employment disciplinary problems, and certain educational attainment should not only be established, but also be enforced (despite the perceived need, or pressure, from the community to hire police officers rapidly). These standards could translate into something as simple as looking closely in the applicants' files for information that is there already—which the NYPD omitted to do in at least 20% of the 400 police officers dismissed or suspended for corruption (Mollen Commission, 1994), and the LAPD failed to do with the 4 out of 14 cases of police officers involved in the Rampart area scandal (LAPD, 2000). It also might mean increasing the processing time to allow investigators to complete background searches, hopefully before the applicants start the police academy, and definitely before they become rookie police officers (as the Mollen Commission discovered, these background checks could be a higher standard to attain [1994]).

Measuring the Prevalence of Police Misconduct

In its attempt to measure the prevalence of career-ending police misconduct, the Kane and White (2009) study relies on official agency data. Kane and White acknowledge that this reliance is a serious limitation of the study. As is the case with any other official data (e.g., arrest rates or crimes known to the police), official records of police officers fired for misconduct could be measuring the tip of the iceberg. Similar to victims of regular street crimes who may be unwilling to report crimes to the police for various reasons—sometimes leading to extremes in which most crimes in some parts of the world are *not* being reported to the police (e.g., Dijk, 1999)—citizens and police officers alike might be not motivated to report (e.g., if the agency has the image that it would not, or could not, do anything about the police misconduct, or if they would be punished or disciplined for participating in the transaction themselves); and they also could be motivated *not* to report police misconduct to the police agency (e.g., if they initiated the corrupt transaction).

A few examples illustrate the point that official agency records tend to be severe underestimates of the true extent of police misconduct. At the time the Knapp Commission reported widespread corruption in the NYPD (1972), which included officers and supervisors reaching the captain level, Cohen's (1972) study of corruption complaints recorded by the NYPD estimated that the number of corruption complaints was less than 1 per 100 police officers annually. As Simpson concluded (1977: 57), Cohen's "figures apparently bear little relationship to reality as indicators of the true level of corrupt activity." Similarly, although the beating of Rodney King resulted in the stipulation of civil liability by the City of Los Angeles, millions of dollars paid to Rodney King, and convictions of two police officers in the federal court, the incident probably never would have ended up in the LAPD's official agency records (despite a serious attempt by Paul King, Rodney King's brother, to report it to the police the next day) were it not for the video recorded by a bystander.

Analogously, for all we know, police officers in the control group could have been engaging in police misconduct less then, about as much as, or more than the police officers in the career-ending group did. The differences reported by Kane and White (2009) could stem from the control group, which includes not only officers who did not engage in serious misconduct, but also officers who did engage in serious misconduct and were caught and resigned before the investigation was completed; or officers who engaged in serious misconduct but were luckier, humbler, smarter, or possessed some other omitted or unobservable characteristic(s) to a larger extent than the officers from the career-ending group (and, thus, the NYPD was less likely to catch, investigate, and fire them). An obvious cautionary tale is that the Kane and White conclusion that less educated police officers were more likely to be associated with official career-ending misconduct—as congruent with common sense as it is—might be interpreted to imply not only that less educated police officers more likely to engage in serious misconduct, but also that the detected police officers were caught because they are not as smart, educated, or otherwise able to avoid detection by the police agency.

Other potential issues with the Kane and White (2009) study notwithstanding, a broad point to be made is that, to try to ameliorate the problems of reliance on official data, future research on this topic should use triangulation—the use of several methodological approaches (e.g., surveys, interviews, and observation)—and thereby obtain more accurate estimates of the nature and extent of police misconduct and, consequently, more reliably compare the characteristics of police officers who did and did not engage in misconduct. For police administrators particularly interested in controlling corruption, this strategy could translate into engaging in proactive investigations, conducting integrity checks, and establishing the early warning systems.

Organizational Approaches

Envision a scenario in which a police department accepts "the most salient policy implications" from Kane and White (2009) and focuses its efforts on the recruitment and selection process. The department would efficiently and consistently reject all applicants with a history of a criminal record and all applicants with disciplinary problems in their prior employment, and it would actively seek applicants with college degrees. Background checks would be thorough, and no one would be hired before their background checks were complete. If these new police officers were then assigned to an NYPD precinct in the late 1960s (see Knapp Commission, 1972), the Philadelphia Police Department in the late 1960s (see Pennsylvania Crime Commission, 1974), the 73rd or 75th precinct in the NYPD in the early 1990s (see Mollen Commission, 1994), or the Rampart Division in the LAPD in the early 1990s (LAPD, 2000), we should have no reasonable expectations that, after 5 years on the job, the overwhelming majority of them—despite their strong initial integrity-related credentials—would remain police officers of high integrity. What went wrong with our carefully selected police applicants? Put differently, why do heightened recruitment standards—even when implemented perfectly and, thus, yielding only high-integrity rookies—fail to make even a dent in the (low) integrity departmental culture?

Many cases (e.g., Christopher Commission, 1991; Knapp Commission, 1972; Mollen Commission, 1994; Pennsylvania Crime Commission, 1974) and research studies (Chevigny, 1995; Fishman, 1978; Sherman, 1978) indicate that the problem with police misconduct typically goes beyond individual officers ("rotten apples") to incorporate problems with police departments ("rotten barrels") or their subunits ("rotten branches"). The organizational and occupational approaches toward police misconduct discuss the effect of police leadership or administration (e.g., Goldstein, 1975; Knapp Commission, 1972; Kutnjak Ivković, 2005; Mollen Commission, 1994), official agency rules (Bittner, 1999; Knapp Commission, 1972; Mollen Commission, 1994), recruitment and selection processes (Burns and Sechrest, 1992; LAPD, 2000; Mollen Commission, 1994), ethics and integrity training (Knapp Commission, 1972; Mollen Commission, 1994), police culture and socialization (Chen, Devery, and Doran, 2003; Klitgaard, 1988; Klockars, Kutnjak Ivković, and Haberfeld, 2004; Skolnick and Fyfe, 1993; Sparrow, Moore, and Kennedy, 1990; Stoddard, 1974; Weisburd, Greenspan, Hamilton, Williams, and Bryant, 2000), supervisors (Burns and Sechrest, 1992; Christopher Commission, 1991; Kappeler, Sluder, and Alpert, 1999; Mollen Commission, 1994; Weisburd et al., 2000), and internal control system (Knapp Commission, 1972; Mollen Commission, 1994; Sherman, 1978) on the extent and nature of police misconduct in a police agency.[1]

Unfortunately, one of the most important issues in this line of research—the effect of the police agency itself and its components on police misconduct—is not addressed by Kane and White (2009). Although this omission was no oversight—it was done on purpose (see Kane and White, 2009: footnote 2)—it nevertheless suggests *caveats* in regard to the findings and ensuing policy implications of the study. This issue might be even more important today than it was decades ago. Although the 1970s commission reports (Knapp Commission, 1972; Pennsylvania Crime Commission, 1974) indicated the widespread presence of grass-variety corruption (Mollen Commission, 1994: E6, 4) at the time, the more recent incidents from the 1990s and the early 2000s (Christopher Commission, 1991; LAPD, 2000; Mollen Commission, 1994) have revealed the emergence of pockets of more aggressive, meat-variety corruption, use of excessive force, and racism. Differences seem to exist not only across police agencies (e.g., Christopher Commission, 1991; Mollen Commission, 1994), but also within smaller subunits

---

1. For example, the Knapp Commission (1972) uncovered systematic, organized, and widespread corruption in the NYPD in the 1970s, which could not have existed without the tacit support for the NYPD administration. As the Knapp Commission (1972: 274) wrote, "the underlying problem is that the climate of the department is inhospitable to attempts to uncover acts of corruption and protective of those who are corrupt." Although the Mollen Commission (1994) did not find corruption to be as widespread and systematic, they reported that "the anti-corruption apparatus was totally ineffective and police department was unable and unwilling to acknowledge and uncover the scope of police corruption … police department's anti-corruption efforts were more committed to avoiding disclosure of corruption than to preventing, detecting, and uprooting it" (1994: E6, 2–3). The Christopher Commission report (1991), written in the aftermath of the Rodney King incident, showed evidence of a police agency in which racism, bias, and use of excessive force were tolerated and concluded that the organizational aspects are crucial: "The LAPD's failure to analyze and act upon these revealing data evidences a significant breakdown in the management and leadership of the department" (Christopher Commission, 1991: iv).

in the same police agency (e.g., Mollen Commission, 1994; LAPD, 2000). For example, the Mollen Commission (1994: E6, 5) reported aggressive forms of corruption and use of excessive force "typically committed by groups of POs assigned to the same command who commit crimes under color of law … [that are] endemic to crime-ridden, narcotics-infested precincts with predominantly minority populations."

Although Kane and White (2009) do not explicitly emphasize and discuss the effect of these organizational factors, their analyses incorporate elements that could be viewed as organizational in nature ("command risk," ranging from "inspector precinct" to "special patrol units;" see Tables 4 and 5). Many of these independent variables are statistically significant in the prediction of all three categories of police misconduct ("crime/serious misconduct," "administrative misconduct," and "drug test failures"). If the data were split across precincts—based on geographic location, nature of the neighborhoods they are serving, the type of leadership, average length of supervisor experience, or ratio of supervisors per line officer—then would some precincts be more likely than others to generate more police officers who were caught and fired for serious misconduct? If the official rules are the same across the NYPD, then what are the conditions that generate different integrity-related outcomes (e.g., overburdened or inexperienced supervisors, drug-infested communities, or a police culture highly tolerant of racism)? Moreover, it is possible that the effects are not the same across various forms of police misconduct (e.g., use of excessive force vs. corruption, corruption vs. perjury) or across different levels of seriousness within the same form (e.g., gratuities vs. kickbacks, tasers vs. deadly force). Learning more about the conditions—be they risk factors or protective factors, as Kane and White call them—and how they relate to police misconduct would contribute substantially to our understanding of the broader problem and specific policy implications.

During the study period, there were inevitable changes in the official rules, organizational culture, community expectations, and opportunities (see the Knapp Commission, 1972 vs. the Mollen Commission, 1994; see also Kane and White, 2009). Time-varying political will, resources, community expectations, and changes in the police department might have translated into different disciplinary environments. Considering subperiods—the few years immediately after the Knapp Commission Report was published (1975–1979, with the greatest emphasis on misconduct control), the post-Knapp period that witnessed the development of crack markets in the early 1980s (1980–1984, with eroding emphasis on misconduct control and recruitment under relaxed standards), and the later period that lead into the 1990s that culminated in the arrest of Michael Dowd (1985–1992, with the weakest emphasis on traditional misconduct control and increased reliance on the drug-use tests)—would have been beneficial.

An Incomplete Tale

Another important point is that multiple, entirely plausible, interpretations exist of the finding that "Black officers were three times more likely to experience career-ending separation for serious misconduct and drug test failures than White officers were" (Kane and White, 2009). An explanation that Kane and White offer is that Black officers were in the minority and that the "tokenism" argument— previously developed by Martin (1980) to describe the female experience—could be applied here as well. Kane and White write the following: "[M]embers of new and growing groups are closely monitored by peers and supervisors in dominant groups, and they are treated with suspicion and a high degree of skepticism about their ability to perform." Thus, because there are so few of them, they are closely monitored and, consequently, misconduct is more likely to be noticed and reported. These arguments lay in the heart of the organizational approaches to police misconduct and emphasize the role of supervisors and peers in shaping the extent and nature of misconduct in a police agency. If combined with other organizational characteristics of their assignment (e.g., precincts with weak supervision, ratio of supervisors to line officers too low, rules enforced rarely and inconsistently; Mollen Commission, 1994), characteristics of the administrative agency adjudication process (e.g., board adjudication; Fyfe, Kane, Grasso, and Ansbro, 1998), and characteristics of the communities they serve (e.g., crime-ridden communities, drug markets; Mollen Commission, 1994), a more accurate picture of why Black officers are more likely to have career-ending misconduct than White officers might emerge.

In this sense, it is intriguing that Kane and White (2009) find the Black and Hispanic versus White race effect for serious/criminal misconduct and drug test/failures but not for administrative misconduct. Several explanations come to mind. First, the administration could have been exercising discretion in the cases involving administrative misconduct. Let us assume that Black officers are more likely than White officers to engage in a career-ending misconduct and that this assumption applies to all three forms of career-ending misconduct, including the administrative one. On the one hand, Kane and White claim that the administration has no choice in the serious/criminal misconduct and drug test failures/refusals. On the other hand, they also claim that the administration has more discretion in cases of administrative misconduct. In the environment in which the community creates strong pressure to hire minority officers and the police agency responds by trying actively to recruit, hire, and retain minority officers, the agency might be *more lenient* toward minority officers than toward White officers. The end result would be no race effect in the multinomial logistic model. It could be that such leniency is not at play. Nonetheless, it is a plausible channel for a result with potentially strong policy implications, and light needs to be shed onto such alternative explanations before policies are shaped and altered.

Second, it could be a consequence of the minority status and its relation to the police culture and "the glass ceiling" (e.g., compared with the more dramatic increases in the other minority group representation, the percentage of black supervisors increased by only 19% in 1990 and

1996; Kane and White, 2009). Police culture has been described as a close-knit community and Black officers, who still represent the minority in the department, might feel isolated from their peers and supervisors. Combined with the effect of tokenism, their conduct is scrutinized closely and is investigated officially at a disproportionately higher rate. Because tokenism and the glass ceiling are closely related, it is to be expected that most supervisors are White. Then, if given a choice, minority officers might be more likely to go through a department trial instead of through the chain of command (which probably provides less transparency), to avoid what they may perceive as racial bias (perpetrated by predominantly White supervisors). Alternatively, Black officers might have less confidence in the police administration to make the right decision, and instead of pleading guilty or going through the chain-of-command procedures, they may prefer a full-fledged administrative hearing, hoping to get a fairer shake. Both of these choices could backfire. As Kane and White (2009) discuss, Fyfe et al. (1998) found that, once the charges and specification were controlled for, "officers who forced departmental trials—regardless of racial background—were significantly more likely than officers who pled guilty (charges and specifications being equal) of being dismissed upon a guilty case disposition."

Third, it also could be the effect of time. Two decades is a long time in the life of an organization, as a comparison of the Knapp Commission (1972) and the Mollen Commission (1994) reports clearly shows. It is possible that the racial effects were stronger and more uniform at the beginning of the period, in the early 1970s, when the percentage of Black police officers was smaller, than they were toward the end of the period in the early 1990s. The changes could be many. The effect of "tokenism" has probably weakened and disappeared toward the end of this period, and the number of Black supervisors has increased; therefore, Black officers might have become less likely to select the administrative hearing, and the police administration might have become more lenient. The combined effect could imply diminishing or disappearing racial differences in the administrative career-ending misconduct. Future research should try to identify factors crucial in this context to guide the development of a more effective policy.

References

Bittner, Egon. 1999. The quasi-military organization of the police. In (Victor E. Kappeler, ed.), *The police and society*, 2nd Edition. Prospect Heights, IL: Waveland Press.

Burns, Pamela and Dale K. Sechrest. 1992. Police corruption: The Miami case. *Criminal Justice and Behavior,* 19: 294–313.

Chen, Janet B.L., Chris Devery, and Sally Doran. 2003. *Fair cop: Learning the art of policing.* Toronto, Canada: University of Toronto Press.

Chevigny, Paul. 1995. *Edge of the knife: Police violence in the Americas.* New York: New Press.

Christopher Commission. 1991. *Report of the Independent Commission on the Los Angeles Police Department.* Los Angeles, CA: Independent Commission on the Los Angeles Police Department.

Cohen, Bernard. 1972. The police internal system of justice in New York City. *Journal of Criminal Law, Criminology, and Police Science,* 63: 54–67.

Delattre, Edwin J. 1989. *Character and cops: Ethics in policing.* Washington, DC: American Enterprise Institute for Public Policy Research.

Dijk, Jan van. 1999. The experience of crime and justice. In (Graeme. Newman, ed.), *Global report on crime and justice*. New York: Oxford University Press.

Fishman, Janet E. 1978. *Measuring police corruption*. New York: John Jay College of Criminal Justice.

Fyfe, James J., Robert Kane, George Grasso, and Michael Ansbro. 1998. *Race, gender, and discipline in the New York Police Department*. Paper presented at the Annual Meeting of the American Society of Criminology, Washington, DC.

Goldstein, Herman. 1975. *Police corruption: A perspective on its nature and control*. Washington, DC: The Police Foundation.

Greisinger, George W., Jeffrey S. Slovak, and Joseph J. Molkup. 1979. *Civil services systems: Their impact on police administration*. Washington, DC: U.S. Department of Justice.

Human Rights Watch. 1998. *Shielded from justice.* New York: Human Rights Watch.

Kane, Robert J. and Michael D. White. 2009. Bad cops: A study of career-ending misconduct among New York City police officers. *Criminology & Public Policy.* This issue.

Kappeler, Victor E., Richard D. Sluder, and Geoffrey P. Alpert. 1999. *Breeding deviant conformity: Police ideology and culture*. Prospect Heights, IL: Waveland Press.

Klitgaard, Robert. 1988. *Controlling corruption*. Berkeley, CA: University of California Press.

Klockars, Carl B., Sanja Kutnjak Ivković, and Maria R. Haberfeld. 2004. *The contours of police integrity.* Thousand Oaks, CA: Sage.

Knapp Commission. 1972. *Knapp Commission report on police corruption*. New York: George Braziller.

Kutnjak Ivković, Sanja. 2005. *Fallen blue knights: Controlling police corruption*. New York: Oxford University Press.

Los Angeles Police Department (LAPD). 2000. *Board of inquiry into the rampart area corruption incident, public report*. Retrieved January 31, 2002 from lapdonline.org/pdf_files/pc/boi_pub.pdf.

Malouff, John M. and Nicola S. Schutte. 1986. Using biographical information to hire the best new police officers: Research findings. *Journal of Police Science and Administration*, 14: 256–267.

Martin, Susan E. 1980. *Breaking and entering: Policewomen on patrol*. Berkeley: University of California Press.

Mollen Commission. 1994. *New York City commission to investigate allegations of police corruption and the anti-corruption procedures of the police department*. New York: Author.

Pennsylvania Crime Commission. 1974. *Report on police corruption and the quality of law enforcement in Philadelphia*. Saint Davids: Author.

Sherman, Lawrence. 1978. *Scandal and reform*. Berkeley: University of California Press.

Simpson, Anthony E. 1977. *The literature of police corruption, Vol. 1: A guide to bibliography and theory*. New York: John Jay Press.

Skolnick, Jerome H. and James J. Fyfe. 1993. *Above the law*. New York: Free Press.

Sparrow, Malcolm K., Mark H. Moore, and David M. Kennedy. 1990. *Beyond 911: A new era for policing*. New York: Basic Books.

Stoddard, Ellwyn R. 1974. A group approach to blue-coat crime. In (Lawrence Sherman, ed.), *Police corruption: A sociological perspective*. Garden City, NJ: Anchor Press.

Weisburd, David, Rosann Greenspan, Edwin E. Hamilton, Hubert Williams, and Kellie A. Bryant. 2000. *Police attitudes toward abuse of authority: Findings from a national survey*. Washington, DC: United States Department of Justice, National Institute of Justice, Research in Brief.

Sanja Kutnjak Ivković, Ph.D., S.J.D., is an associate professor at the Michigan State University, School of Criminal Justice. Her current research interests include police integrity and misconduct, international justice and war crimes, and comparative court systems. Her book that she co-authored with John Hagan, *Reclaiming Justice,* is forthcoming with the Oxford University Press. She has authored *Fallen Blue Knights: Controlling Police Corruption* (Oxford University Press) and *Lay Participation in Criminal Trials* (Austin and Winfield), co-authored *Enhancing Police Integrity* (with Carl B. Klockars and Maria R. Haberfeld; Springer), and co-edited *The Contours of Police Integrity* (with Carl B. Klockars and Maria R. Haberfeld; Sage). Her recent articles appeared in the *Law and Society Review, Journal of Criminal Law and Criminology, Policing: An International Journal of Police Strategies and Management*, *Policing and Society, The Annals of The American Academy of Political and Social Science*, *Law & Policy*, *Law and Social Inquiry, UCLA Journal of International Law and Foreign Affairs, Cornell International Law Journal,* and *Stanford Journal of International Law.* She received her Ph.D. in criminology from University of Delaware in 1995 and S.J.D. from Harvard University, Harvard Law School in 2002.

POLICE MISCONDUCT

Bad cops

Peter K. Manning

Northeastern University

"Corruption," like all fundamental concepts in the social science repertoire, lacks a consensual definition that is either operational, conceptual, or common-sensical. One of my mentors, the late Oxford anthropologist Rodney Needham, urged one to begin a working paper with a close analysis of usage in the Oxford English Dictionary. Unfortunately, such an exercise is beyond the scope of this brief comment. Social scientists long have adopted operational definitions as a way around the quagmire of varying contexts and practices. Although no accepted theory or framework exists for the analysis of corruption, anything referred to as corruption is a product of the everyday practices within an organization— the working rules that make the job possible because they shade and shadow the organization's formal rules (see Moskos, 2008). It is not possible in a work environment to forever adhere strictly to formal rules. Corruption, in simple terms, is a product of an evolved organizational character (Selznick, 1948) and in fact is defined by the organization. It can be accounted for using rhetoric—such as psychological, sociological, or political-economical—but whatever method is used to account for or to explain corruption is determined by the governing administration of the organization at the time. Thus, a corruption scandal typically begins with a denial of wrong doing, slides into a more defined re-specification of causes and consequences, and inevitably ends with a partial admission of culpability (usually entangling those individuals at the bottom as the genesis of the problems revealed). This last stage usually is called "damage control." Abu Ghraib and other recent events provide dramatic illustrations of this natural course of events.

Robert J. Kane and Michael D. White (2009, this issue) argue that, as long as an organization remains able to control the external environment, it can control corruption. This argument makes corruption an organizational problem, not a problem of "attitudes" or individual attributes. However, it is, practically speaking, impossible for an organization to fully control the external environment. An organization might avoid periodic scandals, but it cannot do so indefinitely. The rates, patterns, and cohorts of reported corruption vary from year to year, in part, because of the vulnerability of the organization to penetration. The greater the degree of

Direct correspondence to Peter K. Manning, College of Criminal Justice, Notheastern Univsersity, 400 Church Hall, 360 Huntington Aveune, Boston, MA 02115 (e-mail: pet.manning@neu.edu).

deviance in an organization, the more difficult it is to detect. Generally, in a scandal, one or two people at the top level are transferred or permitted to resign, whereas the bulk of the prosecutions, dismissals, and resignations affect the lower participants. Other than these few generalizations, a radical view would suggest that a comparison of rates across organizations is untenable. Practices assembled to constitute "corruption" are elusive and best studied ethnographically. If one views products of organizational action in summary rule terms—that is, as a consequence of the aggregation of individual decisions—then the attribution of cause or correlation differs from a view of such outcomes as those of constitutive-context bound practices (J. Rawls, 1955; A. Rawls, in press). Studies based on learning theory, those studies applying vignettes, or the case approach used by Klockars, Ivkovich, Harver, and Haberfield (2000) are a considerable distance from anything called police practice. "Bad cops" is a study of attributes, not of practices. It uses official data gathered on members at the bottom of the organization.

Kane and White's (2009) study depends on official data from the New York Police Department (NYPD), uses a flawed dependent variable and sampling practices, does not validate the reported data with other sources (such as interviews and/or ethnographic observations or descriptions), and draws inferences from this material that do not seem to be relevant to the modal U.S. police department (around 20 officers) or even to those departments that serve populations greater than 200,000. It would be misleading to rely on the findings presented by Kane and White to generalize about "corruption" even in large police departments in the United States. These reservations rest not only on flaws in the methodology but also, as I argue below, on the context-based nature of what is labeled "corruption."

Kane and White (2009) gather together several concepts, none of which have been based on detailed examination of practices. Studies of police deviance are not as entirely atheoretical as Kane and White claim. It is also misleading and narrow to restrict relevant studies of "deviance" to those studies focused on the police as an occupation. Relevant organizational studies and the recent work of Peter Moskos (2008) suggest that the context in which deviance is defined renders makes decisions about its sanctioning products of factors outside the immediate context of that behavior. These decisions reflect the locus of organizational power and authority.[1] As Selznick wrote (1948: 27), "In large formal organizations, deviations from the formal system tend to become institutionalized so that 'unwritten laws' and informal associations are established. Institutionalization removes such deviations from the realm of personality differences, transforming them into persistent structural aspects of formal organizations." Sanctions for deviance are methods the top command uses to maintain the power balance of the informal relations over which they have modest control. Such an organizationally grounded perspective means that attributing deviance to categories of people based on identities or personal charac-

---

1.   I can only suggest that the deviance literature, especially the research in the labeling tradition—which was so vibrant in the 1970s—is gone but not forgotten. The fundamental criticisms advanced there concerning official rates and research based on questionnaires, records, and surveys still hold. I would especially recommend Lemert's (1967) concept of "active social control."

teristics is misleading and artifactual. The key question is how these identities are articulated in the informal system.

Consider the sources of data used in the Kane and White (2009) study. Official statistics were gathered, recorded, validated, and stored by the NYPD. In addition, Kane and White gathered materials from records kept by the organization but which had not been assembled for the analysis they carried out. Kane and White are to be applauded for this creative and tedious data-gathering method. Yet even this effort is telling because, as Weber (1947) wrote, the nature of the organization is secrecy, and organizational records either do or do not reflect this nature. Bureaus are organized pockets of secrecy. Since the writings of Gulick and Urwick (1937), it is well known that organizational records are a confirmation of decisions made by more personal means. The informal system of control and association contrasts somewhat with the formally defined administrative system of control. It might be claimed that those individuals at the top rule by exception—bureaucracies strive to routinize all their key procedures to facilitate command control of organizational processes, but they fail repeatedly. Here is the social space in which rulings regarding exceptions take place.

The data were gathered during a period of cycles of reform and reorganization and changes in top command as well as in the high politics of the city (mayor and city council composition). Kane and White (2009) note and describe these changes well. The cohorts of officers inducted into the organization during this period differed in many respects, as the data show. It would have been revealing, given this information, to see a cohort analysis that distinguished the patterns of "deviance" in given cohorts rather than condensing the samples into "good" and "bad" cops. Perhaps the cohorts are more different in composition than the samples of good and bad cops? Although this pattern is suggested in the data on violations by years, it does not reveal the differences within the cohorts who were selected for sanctioning. The term "involuntarily separated" is the basis for the key distinction between the two samples and suggests that organizational pressures to resign worked well. The inclusion of officers who refused to take the drug test—an act "obviously rooted in misconduct," as Kane and White write—surely produces an odd mixture of those labeled "bad cops." It is easy to imagine that officers might refuse on principle or had to decide whether following a principle was worth the job. This subsample surely would be rich domain for future analysis.

The classification scheme for the types of corruption presented is illogical. Although the types were listed from 1 to 8, which implies some logical or even hierarchical rather than indexical order, the scheme does not seem to be based on a single dimension, nor does it reflect "seriousness" or any other obvious ordering. These categories, in behavioral terms are not mutually exclusive (e.g., "drug" violations), include failing to submit to tests as well as possession and sales, but these "misdeeds" are not included under "profit-motivated crimes." In what sense are mere organizational derelictions such as "failure to perform" comparable with legally defined crimes? These "administrative offenses" constitute one third of the reasons given for termination. Also, "failure to perform" is a "garbage" or "sponge" category that the organization

probably uses to dramatize derelictions as "administrative" rather than as criminal. In respect to profit-oriented crimes, which surely have an expressive and symbolic component, how does one read "an apparent goal" from records? As Kane and White (2009) note, the relationships captured by their dependent variable are "complicated." This list of corruption types does not reflect Aristotelian logic: a linear, cumulative, mutually exclusive set of categories arrayed along a single dimension. It reflects labels and categories endemic to the administration of the NYPD. These schemes simply condense reality into a list. The four-category "nominal" variable is not ordinal and, thus, probably has a nonparametric distribution.

This study by Kane and White (2009) is not ethnographic, and it does not claim to be. The issue is whether the classification system used to measure corruption reflects organizational processes and informal power systems. If one were to clarify the nature of corruption within an organization, then (1) one would want to distinguish matters captured in summary rules (aggregated measurement) from matters studied by close ethnographic work—the process by which the instances of "corruption" were identified, labeled, processed, bargained with, and negotiated as to scope and severity. As Kane and White point out, the rates of African Americans, to some degree, are a function of both administrative processes and the reported choices of African Americans to seek adjudication. This observation suggests how organizational practices are the important driving forces in rates of termination. This conclusion is accurate, because it seems that discretion is possible in all cases except drug-related instances in the NYPD and all accounts for the actual processes by which the sanctioning occurs. Unfortunately, the pattern of case attrition by race and violation is not shown. Likely, those individuals with the most serious original charges or allegations were allowed to plead down and resign. Administrative decisions shape the content or category of deviance and the rate of terminations.

(2) One would also want to distinguish and separately analyze categories of corruption by their seriousness as defined by the officers. Some dismissals are pro-forma "bullshit," not respected by the officers themselves because they generally view rule enforcement as capricious, and those incidents thought of as "good work" as part of the territory. (3) One would want to explore the ways in which the incidents came to the attention—how they were reported and by whom—of internal affairs. (4) One would want to distinguish individual violations from repeated instances and distinguish these violations from group offenses—organized crime within the organization itself. (5) It would be useful if one could identify the role of informal social control in the processing of the incident: what kinship, ethnic, class, or other relational factors were part of the allegations, the bargains, or the outcomes represented in these tables.

It is not known what percent of officers in the United States "perform their jobs without engaging in serious deviance." This claim is made by Kane and White (2009), but the study does not reveal supporting evidence regarding NYPD, let alone regarding the range of police organizations in the United States. Unknown levels of deviance are always present in any organization. This statement omits the organizational analog of crime's "dark figure." The notion of "serious deviance" explored here as corruption is administrative—not criminal or civil. It condenses

many behaviors that the organization deems unacceptable and has reported as being processed. The fact that, for more than 100 years, every large city police department in the United States as well as those in Canada, Australia, and England, has experienced several major scandals and cycles of reform, suggests that the processes of interest are not individual attributes. At what point—and why—they become public issues as well as the nature and kinds of organizational practices they reveal is an area rich for investigation. Even a crude assessment of these police scandals would reveal a consistent pattern of the buildup of institutionalized practices that spill into the public attention, often almost accidentally. Many of these instances, such as the willful misrepresentation of crime statistics in Philadelphia and New Orleans—a common organizational practice—are not, strictly speaking, individual corruption. Many acts of deviance (e.g., frame-ups of drug dealers, lying abut the circumstances of observed events, or lying in court) are carried out with the belief that they are fully in the interest of the organization's goals. Acts of violence toward citizens are subject to more media attention, but the question remains of which actions reduce citizen trust of policing more.

The categories "minority," "non-White," and other labels for aggregations (these are not social groups) are profoundly, and deeply, misleading. The fact that rates of "organizational separation" for members of aggregated minority categories taken together show a downward trend, whereas the rates for Black officers remain high despite a decline, is not in itself good news. We have no reason to suspect that any organization treats minorities more justly than others, because the processes of evaluation and assessment are so covert and unavailable short of a civil suit. This decline noted in Kane and White's (2009) data does not, in any way I can conceive of, reflect "better behavior" or a closer representation of the police of "the populations they serve." Whether this should be a principled aim balanced with other aims—such as increased trust, civility in encounters, and fair practices of hiring, promotion, and transfer—is an open question. Whether representational mirroring bears any relationship to the quality of policing is also a moot point. The discussion in this article of "tokens" and "tokenism" is certainly an unfortunate choice of vocabulary. Hughes (1944) was referring to the shaping of sociological and interactional processes by status inequalities within an occupational group and its clients—of which police have none—and evidence for which Kane and White lack.

Is the NYPD in some sense a more democratic organization in which African Americans now have more power, authority, and control of the "informal cultures" in the organization? The distribution of persons by race, rank, and other indicators (Sklansky, 2007) in large police departments shows a thin representation of African Americans at the top level and disproportional representation at middle management where fundamental supervision, punishment, and sanctioning occur. Growth in percentage terms of minorities reported in Kane and White (2009) is also unfortunately misleading. Although the number and percent have increased, the NYPD also has increased radically in size—it is almost a third larger now than 20 years ago—which makes comparisons across time dubious. The absolute size of an organization has effects on power and control. The problems of controlling and sanctioning deviance in an

organization of 40,000 officers differ from the problems in departments smaller than the 10 largest departments in the United States. This is true not only in the span of control sergeants (e.g., the number of officers they supervise and whether they are on shift together), but also in the number of specialized units that are themselves loosely supervised and more democratic in authority patterns. The discovery that proactive investigation assignments do not lead to dismissal is indicative of some organizational processes, as Kane and White note, and is inconsistent with other research (Manning and Redlinger, 1977). Like the claims that broken windows, arrest-based policing leads to a reduction in officially reported crime, this finding cries out for an examination of what these officers do, with whom, and how they do the job. The quality and character of the work done as well as its consequences cannot be measured by officially generated arrest statistics.

But even with the data shown, the advances of African Americans (e.g., reduced levels of termination) compared with other categories did not reveal advances at the same rate. I do not think Kane and White (2009) have any qualitative or observational data to bear on whether the organization's practices and informal culture has changed; therefore, it is unclear what shifts in proportions of groups in ranks mean. Because misconduct is a broad category, and failure or refusal to take a drug test is a simple and easy way to discipline, charging with administrative misconduct is a simple, effective, and "second best" choice.

Kane and White (2009) suggest that those with low self-control demonstrate low self-control more than once. This claim is surely a tautology and not an explanation. The past modes of screening and recruitment in the NYPD have changed according to Kane and White, and it can be expected that high-risk candidates will no longer be hired. The allegation that crime is a function of low self-control and of other self-based frameworks for explaining crime, which include "deterrent" or social learning theories, cannot be proven because individual aggregated data reflect group processes and not individual decisions. Conversely, group processes that affect education, social service, and police practice make individuals subject to stops, questioning, searches, arrest, charge, adjudication, and imprisonment. These facts are incontrovertible. The idea that better screening produces better police officers is a cliché. If this concept means screening out all individuals with a criminal record (which vary widely as to content) as lacking individual merit (which is difficult to define in reference to policing; even the U.S. Supreme Court cannot determine precisely what job-related skills are for individuals applying to the police and fire service), then several social forces collapse, only one of which is "individual choice." This criterion, absent of arrest and of whatever merit it might possess, is confounded by the consistently demonstrated institutionalized differential stop, arrest, and search practices that face minorities in every large city in the United States. Fagan (2009) demonstrated this concept in a striking fashion for African Americans in several sections of New York City. It is likely that approximately 50% of all African Americans less than 24 years of age have arrest records given past data and estimates. Groups based on racial attributes are not treated equally in the United States, and a criminal offense, let alone a felony conviction, disqualifies one for

working even in a fast food restaurant. It is impossible to use such information out of context to make organizational decisions.

Kane and White (2009) mentioned the necessity of panel and longitudinal studies of police careers for future research. What has been shown repeatedly is that social processes—interactions, rewards, learning on the job, and the taken-for-granted—that are the basis for getting along are shaped by (1) the experiences of officers during the first few years on the job, (2) their training officer, and (3) tacit modes of learning. Because these processes replicate the marginalizing interactions of the society at large, it is not news to discover, for example, that in due course, the individuals with valued stereotypic qualities and backgrounds, such as having a relative in the NYPD, are less likely to be terminated. Because it is not evident that the content of police training has changed much since the late 1970s (typically, training is brief, devoted mostly to physical prowess, and deeply anti-intellectual), it is not clear what success in the academy means with respect to a career.

Kane and White (2009) present a preliminary effort to the study of police careers. This area greatly needs exploration. The study of policing, having grown rapidly since approximately the 1960s (Manning, 2004), is still studied as if it were a unique institution rather than a service organization. This article suggests several studies and institutional changes. We lack careful, panel-cohort studies in organizations of different sizes that might follow officers through interviews and records. No proper studies of supervision exist either at the sergeant level or above or at other ranks in middle management. Unlike the other nations in the Anglo-American ambit (New Zealand, Australia, Canada, and Great Britain), the United States lacks a police college as well as an equivalent for graduate training for rising future administrators. The level of training is also comparatively short and uneven across the United States. Training is treated in generalities, and we lack close published studies of police training in the United States. Ironically, the police academy is considered a Siberia, that is, a place to transfer miscreants. No studies of police administrators exist nor do studies of the top command of the 10 or so largest departments. What is known about command, with a few exceptions (e.g., Hunt and Magenau, 1993), is largely anecdotal and autobiographical in the Anglo-American police world. Kane and White (2009) inform us of some consequences of organizational processes and certainly indicate several needed research projects that might reveal what organizational processes these outcomes indicate.

References

Fagan, Jeffery. 2009. *Stops by the NYPD in disadvantaged areas of New York City*. Working paper presented to the International Conference on Intelligence-Led Policing and Security, University of Illinois, Urbana.

Gulick, Luther and Lindal Urwick. 1937. *Papers on the science of administration*. Institute of Public Administration, Columbia University, New York.

Hughes, Everett C. 1944. Dilemmas and contradictions of status. *American Journal of Sociology*, 50: 353–359.

Hunt, Raymond and John M. Magenau. 1993. *Power and the police chief*. Thousand Oaks, CA: Sage.

Kane, Robert J. and Michael D. White. 2009. Bad cops: A study of career-ending misconduct among New York City police officers. *Criminology & Public Policy*. This issue.

Klockars, C., S.K. Ivkovich, W. Harver, and M.R. Haberfield. 2000. The measurement of police integrity. Washington, DC: National Institute of Justice.

Lemert, Edwin. 1967. *Human deviance, social problems, and social control*. Englewood Cliffs, NJ: Prentice-Hall.

Manning, Peter K. 2004. The study of policing. *Police Quarterly*, 8: 23–43.

Manning, Peter K. and Lawrence W. Redlinger. 1977. Working bases for corruption. In (Manning, Peter K. and John Van Maanen, eds.), *Policing: A view from the street*. Santa Monica, CA: Goodyear.

Moskos, Peter. 2008. *Cop in the hood*. Princeton: Princeton University Press.

Rawls, Anne Warfield. 2009. *Journal of Classical Sociology*. In press.

Rawls, John. 1955. Two concepts of rules. *Philosophical Review*, 64: 3–32.

Selznick, Philip. 1948. Foundations for a theory of organizations. *American Sociological Review*, 13: 25–35.

Sklansky, David. 2007. *Democracy and the police*. Stanford, CA: Stanford University Press.

Weber, Max. 1947. Theory of economic and social organization (Talcott Parsons, transl. and ed.). Glencoe, IL: Free Press.

Peter K. Manning (Ph.D. Duke, 1966, MA Oxon, 1982) holds the Elmer V.H. and Eileen M. Brooks Chair in the College of Criminal Justice at Northeastern University, Boston, MA. He has taught at Michigan State, MIT, Oxford, the University at Albany and the University of Michigan, and was a Fellow of the National Institute of Justice, Balliol and Wolfson Colleges, Oxford, the American Bar Foundation, the Rockefeller Villa (Bellagio), and the Centre for Socio-Legal Studies, Wolfson College, Oxford. Listed in *Who's Who in America* and *Who's Who in the World*, he has been awarded many contracts and grants, the Bruce W. Smith and the O.W. Wilson Awards from the Academy of Criminal Justice Sciences, and the Charles Horton Cooley Award from the Michigan Sociological Association. The author and editor of some 19 books, his research interests include democratic policing, uses of information technology, and qualitative methods. His most recent books are *Policing Contingencies* (2003, University of Chicago Press) and *The Technology of Policing* (2008, NYU Press). His current research is focused on information technologies in policing and the transformation of the policing of Ireland following the publication of the Patten Report (1999).

EDITORIAL INTRODUCTION

DETERRENCE AND EXECUTIONS

The impact of the death penalty on murder

John J. Donohue III, *Senior Editor*

Yale Law School

Both history and daily crime sheets underscore a depressing capacity for human violence and inhumanity. Some scholars feel that eliminating capital punishment would be a step toward reducing the toll of human suffering, whereas others feel that retaining the death penalty will prevent some murders at least. Kovandzic, Vieraitis, and Boots (2009, this issue) provide a comprehensive ordinary least-squares (OLS) state panel data assessment of the most recent postmoratorium data available and reach a strong conclusion that the death penalty does not deter murder. This article is an important piece in the complex jigsaw puzzle that will illuminate which factors can deter which crimes under which circumstances.

Commenting on the Kovandzic et al. (2009) article are two scholars who have authored major articles that concern the impact of the death penalty on murder. Richard Berk (2009, this issue) speculates whether the deterrent impact of the death penalty is knowable given current data and methods, whereas Paul Rubin (2009, this issue) argues that "the weight of the evidence as well as the theoretical predictions both argue for deterrence, and econometrically flawed studies such as this article are insufficient to overthrow this presumption." With virtually all positions represented by these three documents, I will discuss three recent new studies that I think address some of Berk's concern and provide strong evidence to support that Kovandzic et al. are right.

The Chicago School and Three Major New Studies

Many Chicago School economists believe that universal answers for deterrence issues flow inexorably from simple price theory. A growing body of evidence contradicts this view, which is often associated with Gary Becker (perhaps unfairly). Predictions about deterrence need to be far more pointed and nuanced than is possible with mere theoretical musings about the slope of demand curves. Undoubtedly, a well-functioning criminal justice system deters massive amounts of crime. If you took the police away—as we have seen when they go on strike or are otherwise

drastically disrupted—then crime can soar.[1] This pattern shows that the Chicago School views capture an important element of truth. But although we see deterrence in one domain, it does not suggest that we will see similar levels of deterrence everywhere.

For example, Lee and McCrary (2009) show that past juvenile offenders are not particularly responsive to the prospect of the dramatically higher sanctions they will face when they reach the age of majority. Specifically, they suggest that a 230% increase in expected sentence length is associated with only a 1.8% reduction in the probability of arrest.[2] In other words, the class of individuals who have been arrested at least once by age 17 shows stunningly little sensitivity to even enormous increases in the probability that they will go to jail and will face a longer sentence if they do. This finding does not refute Becker, but it does suggest that Beckerian deterrence predictions may not apply to at least one class of offenders as they pass to the age of majority.

A second stunning article by Francesco Drago, Roberto Galbiatti, and Pietro Vertova (2009) continues to deepen our understanding of criminal deterrence in two ways. Italy released a huge portion of its prison population on August 1, 2006, but the early release involved the prospect of a penalty kicker; if the early releasees were arrested again, then they would have to repay the time avoided on top of any new sanction. At least in the first 7 months, the threat of the enhanced sentence had a substantial deterrent effect on the average Italian prisoner. Thus, the initial lesson from the Italian experiment vindicates a Chicago School deterrence prediction. But again, deterrent effects are not uniform but are sensitive to context. The Italian prison case revealed that telling a prisoner who has just been let out 3 years early that he or she will have to serve those remaining 3 years tacked on to any additional sentence if caught has a substantial deterrent effect on rearrest for the first 7 months, which becomes somewhat less potent after 1 year (that is all the data we have at present).[3] Will substantial deterrence persist after more than 1 year? We await the results of the follow-up study to answer that question. It also is worth pondering whether a similar deterrent effect would have occurred if a legislative enactment simply had been implemented that all crimes would now carry an extra 3-year sentence. Having a personalized threat may well be more potent than a more uniform and distant threat of punishment.

But a second key finding that emerges from the Italian study conflicts with a wooden Chicago School view of deterrence. The most serious prior criminals—defined as those individuals with an original sentence longer than 69 months—were *not* deterred by the prospect of the enhanced sentence. This point brings us to the death penalty. What Rubin (2009) and many Chicago Schoolers think emerges from theory—an assured deterrent effect—is undermined again by the

---

1. In October 1969, a police strike in Montreal sparked a brief wave of crime and violence. During the strike, banks were robbed, more than 100 shops were looted, and at least 12 fires occurred (Time Magazine, 1969).

2. Although this point estimate is not significantly different from zero, the large standard error means that the largest level of deterrence that emanates from a confidence interval generated by a one-tailed test could be 9.5%.

3. The elasticity of crime with respect to sentence length was estimated at –0.74 for the first 7 months but only –0.45 at 12 months.

best empiricism. The fact that the most serious criminals were *not* deterred by the risk of added jail time in the Italian experiment, and the less serious criminals were deterred, shows that as one marches down the path toward more serious criminal propensities (or toward adolescence, as Lee and McCrary [2009] show), the evidence of deterrence weakens or vanishes.

A final noteworthy study by Zimring, Fagan, and Johnson (in press) that compares the near-identical murder paths of Singapore and Hong Kong despite the latter's renunciation of capital punishment and the former's wildly enthusiastic embrace of it provides yet another indication that Kovandzic et al. (2009) have it right when they conclude that the death penalty has no net deterrent effect in the postmoratorium United States (Zimring et al., 2009).

The Zimring et al. (2009) piece is so powerful because Singapore made Texas look like a piker when it came to using the death penalty; yet the matched-comparison state of Hong Kong experienced the same time path of murder while jettisoning capital punishment. Perhaps someone will show that Singapore's massive jump in executions and subsequent massive decline were tightly calibrated to keep the murder rate moving along a similar path to Hong Kong, or that Hong Kong had some unusually good crime experience when things were going south in Singapore, but the likely answer is what we see from the Italian experience—the most serious criminals are not particularly responsive to threats of higher severity, whether they be a longer sentence or capital punishment.

Berk on Kovandzic, Vieraitis, and Boots

I completely agree with Richard Berk's (2009) assessment that "no credible evidence exists that the death penalty, as implemented in the United States since 1979, has any deterrent value." At present, however, Berk claims "no credible evidence exists to rule out any deterrent effects." He worries that OLS panel data techniques are not powerful enough to answer this enduring issue definitively.

I agree that one panel data study alone can never finally resolve such a contentious question—in part, because the one study itself would need to be thoroughly vetted through continued research. I am more optimistic than Berk (2009), however, that the aggregation of several different types of extremely high-quality studies, which all point in the same direction, can generate more convincing answers to important empirical questions than he seems to think possible. For example, the three studies discussed do not share the same potential weaknesses of OLS approaches. From such combined evidence, a pattern is beginning to emerge that the most serious criminals (and juvenile-offending adolescents) simply are not susceptible to distant threats of heightened punishment, perhaps because they fear they will not be caught, they never think about the consequences of their impulsive acts, or they just do not care. I will be interested to hear if Berk is moved by the Zimring et al. (2009) article to conclude that massive increases in the use of capital punishment followed by an elimination of the death penalty had no discernible effect on murder in Singapore. I would take the Singapore experience as credible

evidence effectively ruling out a deterrent effect of even heavy reliance on executions—at least in Singapore when carried out in the (rather odd) Singaporean way.

Berk (2009) underscores all concerns of trying to tease out causal relationships from observational data, but I think that Kovandzic et al. (2009) do a magnificent job of showing that well-done OLS state panel data analyses on the most current U.S. data simply will not produce evidence of deterrence. This conclusion is an important contribution.

Rubin on Kovandzic, Vieraitis, and Boots

Rubin (2009) also provides some important notes of caution. He observes that the survey evidence amassed by criminologists cannot show that no criminals will be deterred simply because most will not be. He also makes the valid point that perhaps instrumental variable approaches will generate evidence of deterrence when OLS fails. Indeed, the only deterrent effect evidence of capital punishment that survives the Kovandzic et al's (2009) march through the land of OLS estimation are some 2SLS (i.e., two-stage least squares) papers by a handful of authors (which include Rubin and his coauthors).

Although Rubin (2009) suggests that most studies have supported the death penalty, this conclusion is unconvincing because Kovandzic et al. (2009) have laid all of the OLS studies that find deterrence to waste. Kovandzic et al.'s article is now state of the art on OLS estimation of the impact of capital punishment, and its predecessor studies on shorter time frames and with more idiosyncratic specifications simply are off the table now. Those earlier studies no longer can be counted. Progress occurs in science.

Rubin (2009) is correct to note, however, that a hint of life in the deterrence hypothesis can be found in one of the seven Kovandzic et al. (2009) estimates when state-specific time trends are dropped (this evidence is in the model that explains state-year murder rates with a raw number of executions). Against the sea of insignificant results, some scholars may dismiss this one lonely sign of deterrence as spurious, but it is worth thinking about the desirability of including state trends and the "number of executions" specification. Certainly, Kovandzic et al. correct the notion that standard specification tests would call for including the state trends, but such tests are not uniformly dispositive of these vexing specification choices.

However, reasons exist to be skeptical of the number of executions model (recall that even this model only shows signs of deterrence if the state trends are dropped). First, the number of executions model assumes that one would expect a bigger deterrent effect in a huge state that has 2,000 murders and executes 20 murderers (1%) than in a small state that has 10 murders and executes all 10. Certainly on a percentage basis, I would think any deterrence would be far more likely to show up in the latter state, where the message is "all murderers are executed," than in the former (where the message is "it is highly unlikely that you will be executed if you murder someone").

But a second factor provides even greater reason to doubt the isolated "number of executions" model result (when state trends are dropped). Kovandzic et al. (2009) show that this

"number of executions" model also leads to the ostensible finding that more executions lead to a lower *burglary* rate. Because the death penalty for capital murders clearly is not influencing burglaries, this "number of executions" model likely is picking up another effect associated more generally with lower crime.

Rubin (2009) makes other criticisms of Kovandzic et al. (2009) that I find less persuasive. He suggests that including some variables found in Dezhbakhsh, Rubin, and Shepherd (2003) might overturn the Kovandzic et al. findings. I think Kovandzic et al. are correct to steer clear of the arrest and death sentence ratios (as well as the controls for aggravated assault and robbery) that Dezhbakhsh et al. employ, as I have written elsewhere (Donohue and Wolfers, 2009). Rubin's concern that including a measure of racial composition might change Kovandzic et al's results turns out to be unfounded, because adding a control for percent black generates no evidence of deterrence.

Rubin (2009) also questions Kovandzic et al.'s (2009) corrections for serial correlation, but again, I think they follow current best practice in doing so. Rubin's other primary specification critique is that 2SLS approaches are needed, whereas Kovandzic et al. only provide OLS evidence. Kovandzic et al. show that efforts to control for endogeneity in the death penalty arena (they cite Zimmerman, 2006) at times generate clearly biased results. Indeed, Kovandzic et al. as well as Donohue and Wolfers make the case that it is likely that any postmoratorium endogeneity effects would be biased toward finding deterrence. If this view is correct, then it suggests that 2SLS approaches should generate even *less* evidence of deterrence than OLS approaches. Add to that the highly unconvincing instruments that have been offered thus far in the 2SLS studies, and I think Kovandzic et al. were wise to limit their analysis to OLS.

I disagree with Rubin (2009) that "Many selections and remedies [in the Kovandzic et al's (2009) article] are ad hoc and at odds with sound econometric practice." Kovandzic et al. try several different specifications, which may not be preferred, but because some scholars have advocated their use and the results consistently support the no-deterrence finding, it seems reasonable to present the results to clear away any future contention.

Rubin (2009) then states "it would be incredible and a violation of the law of demand if the chance of execution did not deter at least some murders." The point is not well taken, as illustrated in the discussion of the most severe Italian criminals who were not deterred by increased possible sanctions. Even massively increased executions generated no apparent drop in murder in Singapore. But, of course, even if some criminals were deterred by the prospect of the death penalty, some murders might be induced by it, which leads to no net effect. Indeed, Rubin's statement ignores the fundamental asymmetry of potentially benign and malign effects that flow from the death penalty. Specifically, the death penalty only can have a possible useful effect on a small number of individuals—essentially, those individuals who commit murder when they face only life without the possibility of parol—because everyone else already is deterred by lesser sanctions.

For example, in New York—a state with no capital punishment (as of 2004), a large population (19,300,000), and a relatively low murder rate (4.77 per 100,000 people)—we find that 921 murders occurred in 2006. Assuming that 921 roughly equals the number of murderers in New York in 2006, then this figure represents the maximum number of individuals whose behavior could have been changed in a socially acceptable manner by the presence of a death penalty law (at least under a rational actor model). But against these 921 murderers who potentially might have been deterred by capital punishment, approximately 19,299,000 individuals in New York were *not* deterred by the threat of capital punishment (because it was nonexistent and yet they still did not kill). This number is roughly 20,000 times as great as the number of murderers in New York in 2006. If the death penalty has a brutalization or other crime-inducing effect, then we would be concerned that its introduction might have an adverse effect on the 19,299,000 current nonmurderers. If any malign per capita effect were only 1/20,000 as strong as the per capita deterrent effect (potentially influencing only 921 individuals), then the malign effect would offset entirely any deterrent benefit, because it would operate on 20,000 times as many New York citizens. Even if Rubin (2009) was right that some deterrence occurs in the small set of individuals who would murder when confronted by the monumentally severe sentence of life without parole, a tiny per capita crime-inducing effect operating on hundreds of millions could offset or even overwhelm it.[4]

Rubin (2009) concludes his essay with the admonition that "an element of elitism may be present in academic recommendations for abolishing the death penalty, because others will bear the costs." But given the increasingly more powerful evidence that the death penalty neither deters crime in the United States (or Singapore) nor is cost efficient, one must ask, what cost is Rubin thinking of? As I sit in Connecticut where two separate DNA exonerations of convicted murderers finally have released two men after decades in prison, and read the case of what seems to have been an innocent father executed for killing his children (who died in a Texas fire), the case for the death penalty becomes ever more problematic (Labossiere, 2009; Grann, 2009; Herbert, 2009).[5]

---

4. Following a bright line rule that we do not kill people except in the most exigent of circumstances may well restrain harsh and brutal behavior in many different contexts.

5. Cameron Todd Willingham was executed in Texas in 2004 for committing arson that killed his children. According to a recent report, however, no scientific basis determined that the fire was arson. Connecticut resident Kenneth Ireland was sentenced in 1989 to 50 years in prison for the rape and murder of Barbara Pelkey. He was released in August 2009, however, after DNA testing showed he could not have committed the crimes for which he was convicted. After 20 years in prison, Miguel Roman (again from Connecticut) also was released in December 2008 after DNA evidence pointed to the true killer of the woman he was convicted of murdering. In both the Willingham and Roman cases, jail-house snitches were employed to garner convictions when ample reasons were present to be dubious about the prosecutions.

References

Berk, Richard. 2009. Can't tell: Comments on "Does the death penalty save lives?" *Criminology & Public Policy*. This issue.

Dezhbakhsh, Hashem, Paul Rubin, and Joanna M. Shepherd 2003. Does capital punishment have a deterrent effect? New evidence from postmoratorium panel data. *American Law and Economics Review*, 5: 344–376.

Donohue, John J. III and Justin Wolfers. 2009. Estimating the impact of the death penalty on murder. *American Law & Economics Review*, 10.

Drago, Francesco, Roberto Galbiati, and Pietro Vertova. 2009. The deterrent effects of prison: Evidence from a natural experiment. *Journal of Political Economy*, 117: 257–280.

Grann, David. 2009. Trial by fire: Did Texas execute an innocent man? *The New Yorker*, September 7.

Herbert, Bob. 2009. Innocent but dead. *New York Times*. Retrieved September 7, 2009 from nytimes.com/2009/09/01/opinion/01herbert.html?emc=eta1.

Kovandzic, Tomislav V., Lynne M. Vieraitis, and Denise Paquette Boots. 2009. Does the death penalty save lives? New evidence from state panel data, 1977 to 2006. *Criminology & Public Policy*. This issue.

Labossiere, Regine. 2009. Rape, murder charges dismissed against man who served 20 years. *The Hartford Courant.* Retrieved September 7, 2009 from courant.com/community/new-haven/hc-dna-exonerate-0820.artaug20,0,435995.story.

Lee, David and Justin McCrary. 2005. *Crime, punishment, and myopia.* NBER Working Paper No. W11491. Cambridge, MA.

Rubin, Paul H. 2009. Don't scrap the death penalty. *Criminology & Public Policy*. This issue.

Time Magazine. 1969. *Canada: City without cops*. Retrieved September 7, 2009 from time.com/time/magazine/article/0,9171,840236,00.html.

Zimmerman, Paul R. 2006. Estimates of the deterrent effect of alternative execution methods in the United States: 1978–2000. *American Journal of Economics and Sociology,* 65: 909–941.

Zimring, Franklin E., Jeffrey Fagan, and David Johnson. 2009. Executions, deterrence and homicide: A tale of two cities. In press.

John J. Donohue III is the Leighton Homer Surbeck Professor of Law. He is an economist and lawyer who has used large-scale statistical studies to estimate the impact of law and public policy on a wide range of areas from civil rights and employment discrimination law to school funding and crime control. Before joining Yale Law School, he was a chaired professor at both Northwestern Law School and Stanford Law School. He recently published *Employment discrimination: Law and theory.* Among his major articles are: *Uses and abuses of empirical evidence in the death penalty debate* (with Justin Wolfers), *Shooting down the 'more guns, less crime' hypothesis* (with Ian Ayres), and *The impact of legalized abortion on crime* (with Steven Levitt). Donohue is a graduate of Hamilton College, and he received his J.D. from Harvard and a Ph.D. in economics from Yale.

RESEARCH ARTICLE

DETERRENCE AND EXECUTIONS

Does the death penalty save lives?

New evidence from state panel data, 1977 to 2006

Tomislav V. Kovandzic
Lynne M. Vieraitis
Denise Paquette Boots
University of Texas at Dallas

Research Summary

Economists have recently reexamined the "capital punishment deters homicide" thesis using modern econometric methods, with most studies reporting robust deterrent effects. The current study revisits this controversial question using annual state panel data from 1977 to 2006. Employing well-known econometric procedures for panel data analysis, our results provide no empirical support for the argument that the existence or application of the death penalty deters prospective offenders from committing homicide.

Policy Implications

Although policymakers and the public can continue to base support for use of the death penalty on retribution, religion, or other justifications, defending its use based solely on its deterrent effect is contrary to the evidence presented here. At a minimum, policymakers should refrain from justifying its use by claiming that it is a deterrent to homicide and should consider less costly, more effective ways of addressing crime.

Keywords

death penalty, deterrence, homicide, capital punishment, panel data

There may be people on the other side [of the death penalty debate] that rely on older

papers and studies that use outdated statistical techniques or older data, but all of

the modern economic studies in the past decade have found a deterrent effect.

—Joanna Shepherd,
testifying before the Congressional Subcommittee
on Crime, Terrorism, and Homeland Security in 2004

Beginning with the seminal work of Sellin (1959), an extensive body of academic literature has examined the potential deterrent effects of capital punishment on homicide. Sellin's findings that capital punishment had no discernible deterrent effects on homicide, with death penalty (DP) states having murder rates equal to or higher than "matched" abolitionist states (see also Dann, 1935; Savitz, 1958), informed death penalty opinion and policy until the controversial work of Isaac Ehrlich emerged in the *American Economic Review* in 1975. Ehrlich's more sophisticated methodological analysis suggested that each state-sanctioned execution during the 1950s and 1960s "saved eight lives." Moreover, he dismissed the methods employed by Sellin as crude and lacking the necessary scientific rigor to adequately test the complexities of deterrence theory. Ehrlich's findings were well circulated outside of academic circles, where DP advocates effectively transformed his study into a public policy dictate that "proved" the benefits of continuing executions nationally (Fagan, 2005a).

Ehrlich's application of more sophisticated econometric techniques to examine the deterrent effects of the DP was a clear advancement over previous work. Despite these improvements, however, Ehrlich's (1975) study was criticized as suffering from serious empirical infirmities, and as a consequence, its conclusions about the powerful deterrent effects of capital punishment on homicide were later deemed unjustified (Baldus and Cole, 1975; Blumstein, Cohen, and Nagin, 1978; Bowers and Pierce, 1975). Numerous academic papers during the next two decades continued to investigate the potential deterrent effect of capital punishment on homicide, with most criminological studies showing no deterrent effect or even citing a brutalization effect, whereby homicides increased as an unintended consequence of state executions (Cochran and Chamlin, 2000; Lamperti, 2008).

The death penalty debate was ignited once again with the 2003 publication of Dezhbakhsh, Rubin, and Shepherd's study on the deterrent effect of capital punishment. Using county-level panel data for the post-Gregg era, they estimated that 18 lives were saved each year for each execution. These findings of a strong deterrent effect of the death penalty prompted numerous empirical economists to reexamine the DP efficacy hypothesis using modern econometric methods for panel data, with several studies reporting robust deterrent effects. Yet again, this newest generation of economic deterrence studies has received significant attention from the press, DP advocates, and policymakers who are eager to justify punitive crime-control measures such as the DP (Fagan, 2005b, 2006). Although research conducted by criminologists and some econo-

mists has consistently found little or no support for the deterrent effect of the DP on homicide, empirical economists relying heavily on the Beckerian model of crime have largely ignored or summarily dismissed these studies as lacking appropriate methodological rigor.[1] Criminologists and their research are again notably absent from the capital punishment debate.

The current study revisits the controversial question of whether the DP exerts a deterrent effect on the homicide rate using annual state panel data from 1977 to 2006. This article employs many of the same econometric "bells and whistles" used in recent economic papers on the DP, while substantively contributing to the literature regarding the deterrence hypothesis debate as it (1) remedies statistical problems found in several recent DP studies reporting robust deterrent effects; (2) controls for a larger number of potential confounding factors that are theoretically grounded, including several crime policy variables (e.g., three-strikes laws [3X] and right-to-carry concealed handgun laws) and historical events (e.g., U.S. imprisonment binge and crack-cocaine epidemic of the 1980s) that have been linked with cross-temporal changes in homicide rates in the post-moratorium era; and (3) extends the analysis to include additional years (beyond 2000) not covered in recent state panel DP papers. The following section begins with a review of recent economic papers on the DP. We then describe our data and methods and present our results. In the final section, we interpret our results with reference to criminological research on rational choice and offender decision making and consider the policy implications of our findings.

Background

During the last 10 years, an upsurge has occurred in the number of empirical studies, mostly by economists, estimating the average deterrent effect of the DP on homicide rates across states with capital punishment. These studies have primarily relied on annual state- or county-level panel data using fixed-effects models and have operated within the ordinary least-squares (OLS) estimator framework.[2] Although some panel research has extended the study period prior to the death penalty moratorium that began with *Furman v. Georgia* in 1972 (e.g., Dezhbakhsh and Shepherd, 2006; Katz, Levitt, and Shustorovich, 2003), most have focused on within-state (or -county) changes in the overall homicide rates after the reinstatement of capital punishment in *Gregg v. Georgia* (1976; beginning in 1977 or later). The main differences among the fixed-effects panel studies are the ways in which the authors have conceptualized and operationalized

---

1. In simple terms, Becker's (1967) principles advance a strongly prorationality position whereby decision making is propelled by cost-driven calculus, such that offenders commit crimes because the potential benefits outweigh the potential risks. It should be noted, however, that other economists have questioned these prodeterrence studies after finding contradictory results and that not all economists endorse Beckerian principles.

2. A few of these studies employ other quasi-experimental designs to analyze the effects of governor- or court-imposed moratoria (e.g., Cloninger and Marchesini, 2006) or the extent to which the effects of execution risk are contingent on newspaper publicity surrounding executions (Stolzenberg and D'Alessio, 2004). Because these studies focus on potential deterrent effects operating in a single DP jurisdiction, as well as for ease of presentation, we did not include them in our review.

execution risk. Given these differences, and for ease of presentation, we organized our review of the latest DP deterrence research based on the measures of execution risk used in each study (i.e., presence of the DP, probability of execution, and frequency of execution). In the next section, we discuss the methodological shortcomings of studies employing econometric methods for panel data and how these problems are mitigated in the current study.

The findings of the latest DP deterrence studies using state- or county-level panel data are summarized in Table 1.[3] The table includes the time period covered, unit of analysis, measure used to denote activity status of DP statute or execution risk, and the results obtained for these measures. Comprehensive reviews of the latest DP deterrence studies can also be found in Donohue and Wolfers, (2005), Fagan (2006), Shepherd (2005), and Yang and Lester (2008).

Presence of the Death Penalty

Of the 10 studies published since 2000, 6 examined whether the mere presence (or absence, because of a moratorium or the law being abolished) of the DP was a deterrent to homicide by entering a binary dummy variable into the regression model that took on the value of 1 if the DP was legal in the state and 0 otherwise (Dezhbakhsh et al., 2003; Dezhbakhsh and Shepherd, 2006; Donohue and Wolfers, 2005; Ekelund, Jackson, Ressler, and Tollison, 2006; Mocan and Gittings, 2003; Zimmerman, 2006).[4] The dummy variable approach implicitly assumes that the deterrent effects of the DP are unrelated to the probability of execution; rather, the mere existence of capital punishment is assumed to exert a deterrent effect that is not systematically stronger in years with higher actual probabilities of execution.

With the exception of Ekelund et al. (2006), the preponderance of the evidence indicates that the presence of a DP statute was associated with lower homicide rates, although the negative coefficients for the DP dummy variable reported by Donohue and Wolfers (2005) were not significant at conventional significance levels. Specifically, Mocan and Gittings (2003) reported that the presence of the DP reduced the annual number of homicides by 64, whereas Zimmerman (2006) concluded that deterrent effects attributed to the presence of the DP were similar for all five methods of execution. The most notable study to use the dummy variable approach, conducted by Dezhbakhsh and Shepherd (2006), treated the U.S. Supreme Court's 1972 decision imposing a moratorium on the DP as a "judicial experiment" by coding states a 1 for each year in which the moratorium was in effect and 0 otherwise. In all specifications (see their Table 8), the coefficient on the DP dummy variable was significant and positive, which indicates that stopping executions increased the homicide rate or that reinstating the DP reduced the homicide rate. Conversely, Ekelund et al. (2006) reported results across specifications that, with a single exception, were statistically significant and positive, which suggests that the presence of an active DP law actually increased homicide during the 1995 to 2000 period.

---

3.   The studies in Table 1 are limited to those published in 2000 or later.

4.   Dezhbakhsh and Shepherd (2006) switch the coding so that states with an inactive DP law are coded 1 and 0 otherwise.

Probability of Execution

With the exception of Dezhbakhsh and Shepherd (2006), all studies listed in Table 1 included some measure of the probability that an offender would be executed. In general, the probability of execution was operationalized as (1) the ratio of the number of executions to the number of homicides (Zimmerman, 2006); (2) the ratio of the number of executions to the number of inmates on death row (Shepherd, 2004); (3) the ratio of the number of executions to the number of offenders sentenced to death (Dezhbakhsh et al., 2003; Mocan and Gittings, 2003; Shepherd, 2005; Zimmerman, 2004); (4) the ratio of the number of executions to the number of prisoners (Donohue and Wolfers, 2005; Katz et al., 2003); and (5) the ratio of the number of executions to population (Donohue and Wolfers, 2005). Again, the main difference occurs in the denominator, where scholars have largely disagreed on the total number of possible outcomes potential murderers are likely to consider when calculating these risks.

Most researchers have used lags ranging from approximately 1 to 6 years in the denominator based on when they expect an execution to impact the homicide rate. The amount of time each variable is lagged depends on the scholar's estimation of the criminal calculus and/or the processing of an offender through the criminal justice system from arrest to execution. For example, Mocan and Gittings (2003) and Shepherd (2005) use a 6-year lag in their execution risk measures. The justification for the use of a 6-year lag is based on Bedau's (1997: 15) estimation that it takes an average of 6 years for an offender to be executed after being sentenced to death (an estimation based on data in the Bureau of Justice Statistics [BJS] report, *Capital Punishment, 1994)*. From a deterrence perspective, potential murderers would conduct a cost–benefit analysis based on the numbers of offenders sentenced to death 6 years before, rather than on current-year sentences. Thus, if offenders are influenced by the probability they will be sentenced and executed, then they would calculate their risk and likelihood based on current-year executions of death row inmates who had been sentenced 6 years earlier.

Other scholars have used a shorter time period under the assumption that offenders will base their decisions on whether to commit homicide on what is currently or recently happened to friends or acquaintances (Donohue and Wolfers, 2005; Shepherd, 2004, 2005). For example, some scholars have defined the probability of execution using current-year death sentences in the denominator of the ratio variable (e.g., Zimmerman, 2004), arguing that prospective murderers are unlikely to compute actual probabilities for cohorts of convicted murderers because doing so would be extremely costly for the potential murderer (Shepherd, 2004; Zimmerman, 2004). These scholars have maintained that potential murderers are likely to form expectations based on a "cheaper informational proxy," such as the current going rate at which convicted murderers are sentenced to death row and executed (Shepherd, 2004: 297).

Regardless of how the probability of execution is measured, studies generally report a negative association between execution risk and the homicide rate, but statistical significance has varied. Katz et al. (2003) reported that the coefficients for the execution rates entered in their regression models were extremely sensitive to model specification and were sometimes positive

T A B L E   1

Summary of State-Level Panel Studies on the Impact of the Death Penalty on Homicide Rates

Study	Time Period	Sample	Measures	Impact of Death Penalty[a]
Dezhbakhsh, Rubin, and Shepherd (2003)	1977–1996	3,054 counties	*Statute dummy*	HOMICIDE – **
			Sentencing rate lagged 2 years: death sentences $(t + 2 + t + 1 + t + t − 1 + t − 2 + t − 3)$/arrests $(t − 4 + t − 5 + t − 6 + t − 7 + t − 8 + t − 9)$	Homicide –
			Execution rate lagged 2 years: death sentences $(t + 2 + t + 1 + t + t − 1 + t − 2 + t − 3)$/executions $(t − 4 + t − 5 + t − 6 + t − 7 + t − 8 + t − 9)$	HOMICIDE – **
Katz, Levitt, and Shustorovich (2003)	1950–1990	50 states	*Execution rate:* executions per 1,000 prisoners	Homicide –
			Execution rate lagged 1 year	Homicide +
			Execution rate lagged 2 years	Homicide +
			Execution rate lagged 3 years	Homicide +
Mocan and Gittings (2003)	1977–1997	50 states	*Statute dummy lagged 1 year*	HOMICIDE – ***
			Sentencing rate lagged 1 year: death sent/ murder arrests 2 years prior	Homicide +
			Execution rate lagged 1 year: executions previous & current year/ death sent 6 years prior	HOMICIDE – **
			Execution rate 2 lagged 1 year: executions first 3 quarters of current year and last quarter of previous year/number persons sent to death 6 years prior	Homicide –
Zimmerman (2004)	1978–1997	50 states	*Sentencing rate:* death sentences/murder arrests	Homicide –
			Execution rate: executions/death sentences	HOMICIDE – **
Shepherd (2004)[b]	1977–1999	50 states	*Sentencing rate:* moving average number of death sent in current & previous 11 months/12 month moving average of number of murders	HOMICIDE – **
			Execution rate: 12 month moving average of number of executions/12 months moving average of number on death row	HOMICIDE – **
			Number of death sentences	Homicide +
			Number of executions	HOMICIDE – **

Study	Years	Sample	Variable	Result
Shepherd (2005)	1977–1996	50 states	Sentencing rate 1: number of death sentences/number of arrests for murder-2	HOMICIDE + (13 states)[c]
			Sentencing rate 2: number of death sentences +2/no of arrests for murder	HOMICIDE − (6 states)
			Execution rate 1: number of executions/number of death sentences − 6	Homicide + (8 states)
			Execution rate 2: number of executions + 6/number of death sentences	
			Execution rate 3: executions $t + 2, t + 1, t, t − 1, t − 2,$ $t − 3$/death sentences $t − 4, t − 5, t − 6, t − 7, t − 8, t − 9$	
Donohue and Wolfers (2005)	1960–2000	50 states	Statute dummy	Homicide −
	1934–2000	50 states	Active law: ≥ 1 execution in previous decade	Homicide −
	1960–2000	50 states	Inactive law: no executions in previous decade	Homicide −
	1978–1997	50 states	Execution rate: executions/1,000 prisoners	Homicide +
			Number of executions	HOMICIDE −***
			Execution rate: executions/100,000 pop.	Homicide −
			Execution rate 2: executions/1,000 prisoners	Homicide −
			Execution rate 3: executions/homicide −1	Homicide −
			Execution rate: executions/death sent lagged 1 year	Homicide +
Dezhbakhsh and Shepherd (2006)	1960–2000	50 states & DC	Number of executions	HOMICIDE −***
			Number of executions lagged 1 year	HOMICIDE −***
			State moratorium dummy	HOMICIDE + ***
Ekelund, Jackson, Ressler, and Tollison (2006)	1995–2000	50 states & DC	Statute dummy	HOMICIDE+***
			Number of executions lagged 1 year	HOMICIDE−***
Zimmerman (2006)	1978–2000	50 states	Statute dummy	HOMICIDE −**
			Execution rate: electrocutions/homicides	HOMICIDE −**

[a] Capitalization means the coefficient is significant.

[b] Results reported for total homicide only.

[c] Shepherd transformed the statistically significant results from all models into each state's increase or decrease in the number of murders after one execution.

* $p = .10.$ ** $p = .05.$ *** $p = .01.$

and sometimes negative. Donohue and Wolfers (2005) generally found no statistically significant association between execution risk and homicide rates, whereas Dezhbakhsh et al. (2003), Mocan and Gittings (2003), Zimmerman (2004, 2006), and Shepherd (2004) reported robust deterrent effects. In Zimmerman's (2006) study, however, these effects were significant only for executions by electrocution. None of the other four methods (i.e., lethal gas, lethal injection, hanging, or firing squad) had a significant impact on homicide rates. Finally, Shepherd (2005) found a "threshold effect," meaning that states that executed more than nine persons during the sample period executions observed lower homicide rates, whereas states that conducted fewer executions had higher homicide rates.

Frequency of Execution

The most widely used measure of execution risk in DP deterrence studies has been the frequency of executions (e.g., Dezhbakhsh and Shepherd, 2006; Ekelund et al., 2006; Shepherd, 2004). This conceptualization of deterrence suggests that optimal deterrence is most likely to be realized by simply "reminding" prospective murderers of the state's willingness to use capital punishment to deter homicide (Kleck, 1979: 896). Thus, regardless of whether prospective murderers are inclined to or capable of calculating the probability of being executed for murder, such persons might still be deterred if increases in executions cause increases in their perceptions of execution risk (presumably through "publicity effects"). Donohue and Wolfers (2005) were critical of the frequency of execution measure because (1) it has the net effect of giving high-execution states, such as Texas and Virginia, greater weight in the homicide regression models and (2) it implies that the effect of an additional execution will vary across DP states depending on the size of the population.[5] Dezhbakhsh and Rubin (2007: 17) responded to the criticism levied by Donohue and Wolfers by arguing

> an execution in a densely populated state with more crimes, more criminals, and more potential criminals has a stronger deterrent effect, in terms of the number of lives saved, than an execution in a sparsely populated state with few crimes and few potential criminals. So dividing the number of executions by population makes no sense.

Although the points raised by Donohue and Wolfers call into serious question the theoretical underpinnings used to justify the frequency of executions as a measure of execution risk, we cannot rule it out as one of many possible scenarios through which executions may have the effect of deterring homicide offenders. Indeed, all four studies employing frequency of executions

---

5. For example, Dezhbakhsh and Shepherd (2006) estimate the effect of an execution on the homicide rate to be –0.145, which implies that each execution in Texas reduces the annual number of homicides by roughly 20, whereas in Delaware, it reduces the annual number of homicides by almost 1.

as a measure of execution risk found strong support for the DP deterrence-efficacy hypothesis. These results suggest that executions might exert a unique deterrent effect on homicide rates even in years when the actual probability of execution for murder in the same state is less than in previous years or greater than in other DP states (Dezhbakhsh and Shepherd, 2006; Donohue and Wolfers, 2005; Ekelund et al., 2006; Shepherd, 2004).

In sum, although most scholars studying the deterrent effects of the DP have agreed that deterrence depends more heavily on the actual risk of execution rather than on the mere existence of the DP (e.g., Dezhbakhsh and Rubin, 2007; Mocan and Gittings, 2003), they have differed on which factors prospective murderers consider when calculating such risks. Given the lack of reliable information on how prospective murderers assess the risk of execution, if at all, it is not surprising that there is no theoretical or empirical consensus on how best to measure execution risk. More importantly, however, DP scholars have necessarily assumed that any such measure of actual execution risk would have a positive effect on average perceptions of execution risk among prospective murderers. Research by Kleck, Sever, Li, and Gertz (2005) suggests, however, that the perceived risk of punishment has little or no relationship to the actual risk of punishment, and this finding may apply specifically to the risk of execution.

Data and Statistical Methods

Similar to recent economic papers on capital punishment, we reexamine the "DP deters homicide thesis" using annual, state panel data. Because we are solely interested in assessing potential deterrent effects of capital punishment in the post-Gregg era, we begin our study period in 1977 but extend the study period used in recent studies from 2000 to 2006.[6] The primary advantage of the panel design, as opposed to the more commonly used time-series design (e.g., national time-series studies) in earlier DP deterrence research, is that it provides a comparison group by treating non-DP states as a control group for DP states (Campbell and Stanley, 1963). We are not, however, asserting that non-DP states represent a control group in the strict sense of the term, as this would imply the DP is a "natural experiment." The defining feature of a true natural experiment is that assignment of treatment conditions occur in an "as if" random fashion (Dunning, 2005). Because it is unlikely that the decision to enact, abolish, halt, or apply the DP in the post-Gregg era occurs independently of other sociopolitical forces operating in DP states, we do not believe a credible claim can be made that DP and non-DP states are "as if" randomly assigned. As a result, we do not assume "pretreatment" equivalence between DP and non-DP states; rather, we follow the standard solution in nonexperimental research (and recent DP deterrence papers) of measuring and statistically controlling for as many potential confounding factors—that is, correlates of the presence of the DP and risk of execution that could influence homicide rates—as possible. We also follow common practice in state panel studies of the DP (and crime-control initiatives in general) by including state fixed effects, year fixed effects, and state-specific time trends in the homicide specifications to minimize potential

6.   None of the latest DP panel studies of which we are aware uses data that extends past 2000 (see Table 1).

endogeneity problems related to omitted variable bias. A detailed discussion of how these proxy variables minimize the effects of omitted variable bias is provided in the next section.

Death Penalty Measures

As discussed, deterrence theory provides multiple pathways by which the DP could serve as a deterrent to potential murderers. Given the nature of the study and the lack of a general consensus on how prospective murderers might form these expectations (assuming they will do so at all), we felt it was important to include all of them in the current study. The measures used to capture both the presence of the death penalty and execution risk are as follows:

Death Penalty Law Status Variables:
1.   DP law dummy variable (year t)
2.   DP law dummy variable (year $t-1$)

Frequency of Executions:
3.   Number of executions (year t)
4.   Number of executions (year $t-1$)

Probability of Execution Measures:
5.   Executions (year t) per 1,000 prisoners (year t)
6.   Executions (year t)/death sentences (year $t-1$)
7.   Executions (year t)/death sentences (year $t-6$)
8.   Executions (year t) per 100,000 state population (year t)
9.   Executions (year t)/homicides (year $t-1$)

Data regarding the legal status of the DP through 2000 were obtained from Dezhbakhsh and Shepherd (2006: 513). Using sources cited in Dezhbakhsh and Shepherd, we collected additional information on the legal status of the DP through 2006. Year-end data on the number of prisoners currently on death row; those receiving a death sentence or executed in the current year; or those persons removed from death row because of a sentence vacation/commutation, death from natural causes, suicide, escape, or drug overdose from 1977 to 2005, came from "Capital Punishment in the United States, 1973–2005" (Bureau of Justice Statistics, 2005).

Year-end statistical tables and data from 2006 were downloaded from the Bureau of Justice Statistics (2007) Web site as well.[7]

Homicide Rates

It has been suggested by some that, because the DP can only be applied to capital murders, the most appropriate dependent variable in a DP deterrence study is the rate of capital murders (e.g., Bailey, 1983; Fagan, Zimring, and Geller, 2006; Peterson and Bailey, 1991; Sellin, 1959; Van den Haag, 1969). We maintain that an equally valid argument can be made for the use of the total homicide rate to test the DP efficacy hypothesis. Drawing on Van den Haag's (1969) conceptualization of deterrence, Kleck (1979) argued that the deterrent effects of the DP need not be limited to prospective offenders engaging consciously in risk–benefit calculations but to all homicides, as "the cognitive link in potential offenders' minds may be between the ultimate legal sanction, death, and the act of homicide rather than any particular arbitrary legal subtype of homicide." (1979: 890). Kleck's application of Van den Haag's preconscious deterrence theory to the DP provides a theoretical rationale for broadening the search for potential deterrent effects by including both death and non-death-eligible homicides in the homicide rate measure (see Shepherd, 2004, for empirical support).

Homicide data were obtained from the Federal Bureau of Investigation's (FBI's) Uniform Crime Reporting (UCR) Program, published as *Crime in the United States*. UCR homicide data from 1977 to 2006 are available on-line at the BJS Web site (ojp.usdoj.gov/bjs/dtd.htm). We rely on the FBI's UCR homicide measure—as opposed to homicide data based on death certificates collected as part of the National Vital Statistics System by the National Center for Health Statistics (NCHS)—because the latter are available only through 2005.[8]

---

7. Analyses that measure execution risk based on the number of death sentences issued in the previous year or 6 years prior cover the period 1978 to 2006 and 1984 to 2006, respectively. The rationale for excluding earlier years was that few criminals were sentenced to death during the 4-year hiatus (1972–1976) on capital punishment after the Furman v. Georgia (1972) decision. As a result, measures of execution risk calculated using death sentences meted out during the years of the ban would be undefined (because of the zero denominators), and it is impossible to know how prospective murderers during the years 1978 to 1982, for example (assuming death sentences meted out in the previous 6 years is the correct denominator), would have calculated their risk of execution. Even after excluding these years, the measures of execution risk remained undefined in many instances because no death sentences were issued by the state in the previous year (or 6 years earlier). To avoid losing these state/years in the analysis, undefined observations were assigned a score of 0. Coefficient estimates for the ratio variables were qualitatively similar when treating undefined observations as missing data.

8. After the data analysis was completed, we obtained homicide data based on death certificate data for 2006 using the Centers for Disease Control and Prevention's WISQARS interactive database system. Although not shown, the sign and size of the coefficients obtained for the execution risk measures were largely similar when substituting the FBI's UCR homicide measure for the NCHS homicide measure.

Specific Control Variables

Crime policy initiatives and the crack epidemic. As discussed, most of the latest DP papers failed to account for other important crime-control initiatives or important historical events that occurred in the post-moratorium era. The passage of "three strikes and you're out laws," for example, have been linked with homicide increases (Kovandzic, Sloan, and Vieraitis, 2002; Marvell and Moody, 2001) and decreases (Shepherd, 2002).[9] In addition, a large number of academic studies have examined the potential deterrent impact of right-to-carry concealed handgun (RTC) laws on homicide rates, with mixed results. Although Lott and Mustard (1997) and Lott (2000) reported robust evidence of deterrence, several researchers who reanalyzed (and extended) their data (e.g., Ayres and Donohue, 2003, 2009a, 2009b) concluded that "the statistical evidence that these [RTC laws] have reduced crime is limited, sporadic, and extraordinarily fragile" (Ayres and Donohue, 2003: 1201). In any event, several published studies support the RTC law-efficacy hypothesis, and we err on the side of caution by including an RTC law variable as a regressor to avoid a potential model underfitting problem. Both laws are represented with a binary dummy variable scored "1" starting the full first year after a law went into effect and "0" otherwise. Dates for the passage of 3X laws were obtained from Marvell and Moody (2001). Dates of passage for RTC laws were obtained through statutory research conducted by Marvell (2001) and the senior author.

We also control for the prevalence of crack cocaine using an index created by Fryer, Heaton, Levitt, and Murphy (2005). The crack index is composed of various indirect proxies of crack prevalence, including cocaine arrests, cocaine-related emergency room visits, cocaine-induced drug deaths, crack mentions in newspapers, and Drug Enforcement Administration drug busts. Unfortunately, the crack index variable is only available through 2000. Rather than shorten the study period by 6 years, we only enter the crack index variable in separate estimations when examining the robustness of the baseline homicide specifications. Data for the crack prevalence measure were obtained from Roland Fryer's Web site at post.economics.harvard.edu/faculty/fryer/fryer/.html.

Socioeconomic control variables. Socioeconomic variables included in the homicide specifications are those commonly used in recent DP papers and in macrolevel studies of homicide in general. Specifically, we control for the percent of the civilian labor force unemployed; the total employment rate; real per-capita income (divided by the Consumer Price Index); percent of the population living below the poverty line; percent of the population residing in metropolitan areas; percent of the population with a bachelor's degree or higher; per-capita beer consumption (measured in gallons); and the percent of the population ages 15 to 24, 25 to 34, and 35 to 44 years. Poverty data were obtained from the Bureau of the Census Web site at census.gov/hhes/www/poverty. The data on state-level unemployment were taken from the Bureau of Labor

9.   Between 1993 and 1996, 25 states and the federal government enacted 3X laws (Austin and Irwin, 2001). Studies that found a positive association between 3X laws and homicide speculate that felons who face lengthy prison terms after conviction for a third strike may decide to kill victims, witnesses, or police officers during an attack that would otherwise be nonlethal to reduce the chance of apprehension.

Statistics Web site at bls.gov/sae/home. Data on personal income and real welfare payments were obtained from the Bureau of Economic Analysis Web site at bea.doc.gov/bea/regional/reis/. The percent of the population with college degrees or higher and residing in metropolitan areas are linear interpolations of decennial census data, as reported in various editions of the *Statistical Abstracts of the United States*. Data on beer consumption were obtained from the Beer Institute Web site at beerinstitute.org. The age group data were obtained directly from the U.S. Bureau of the Census on computer disk.

Deterrence measures. Deterrence measures include the number of police officers per 100,000 population, the state incarceration rate (again, per 100,000 population), and the prison death rate.[10] The latter measure is used to proxy for the quality of life in prisons, which Katz et al. (2003) argued and demonstrated to be a deterrent to criminals. The variable is defined as the number of prisoners who die in prison from all sources (except executions) per 1,000 state prisoners. The data on the total number of police (including civilians) were from the Public Employment series prepared by the Bureau of the Census. The data on the number of prisoners (sentenced to prison more than a year in custody as of December 31) were obtained directly from the BJS Web site at ojp.usdoj.gov/bjs. Prison death data from 1977 to 2000 were obtained from Justin Wolfers' personal Web site at bpp.wharton.upenn.edu/jwolfers/DeathPenalty.shtml. The data from 2001 to 2006 were obtained from the BJS's "Deaths in Custody Reporting Program" (Bureau of Justice Statistics, 2008).

Statistical Methods for Panel Data

We follow conventional strategies for panel studies of crime and estimate a fixed-effects model. The fixed-effects model requires adding a dummy variable for each state and year (except the first to avoid dummy variable trap). The state (cross-sectional) fixed effects control for time-invariant unobserved factors that influence homicide rates in a particular state.[11] The year fixed

---

10. The data are for the end of the year, and we estimate the prison population during the year by averaging the current and prior year numbers. We follow the conventional practice of lagging the police and prison measures by 1 year to mitigate potential simultaneity bias.

11. The Hausman specification tests strongly rejected the notion of no systematic differences between the fixed-effects and random-effects models.

effects control for unobserved factors that are common to all states in a particular year.[12] Because the analysis includes both state and year fixed effects, the parameter estimates for all explanatory variables are based solely on within-state variation. We also control for state-specific time trends by including a separate linear trend variable for each state. The inclusion of state-specific time trends has become standard practice in panel studies of crime (e.g., Ayres and Donohue, 2003; Donohue and Wolfers, 2005; Marvell and Moody, 1996; Mocan and Gittings, 2003). The state-specific linear time trends control for unobserved, slow-moving, sociodemographic factors that affect the time-series behavior of the homicide rate in each state and differ from the nationwide trends captured by the year fixed effects.

The next issue for the homicide rate model regards choice of functional form. The most common practice in the recent DP literature and the procedure followed here is to assume a linear functional relationship between the risk of execution and homicide rates. Although the preference in state panel studies of crime is to use a log-log model, this would force us to address the problem of having a large number of 0-valued observations for the execution risk measures that cannot be logged. It is common to add some small amount (usually a 1) to 0 values so that a logarithmic functional form may be used. However, this procedure is inappropriate where, as here, there are a large number of 0s (Wooldridge, 2005: 185). Because we have no strong theoretical basis for choosing one functional form over another, we also present alternative estimates using a log-level model in which only the dependent variable (i.e., the homicide rate) is logged. The log-level model is intuitively appealing, especially for those who champion the economic model of crime, because it implies that as execution risk increases, the deterrent effects of the DP on the homicide rate will accelerate.

We also follow recent convention in panel data analysis of assuming "clustered errors" and compute heteroskedasticity-autocorrelation robust standard errors clustered at the state level. The benefit of using cluster robust standard errors is that they allow for general forms of heteroskedasticity as well as for arbitrary serial correlation within a given state (Wooldridge, 2001). Failing to account for the presence of clustered errors produces biased estimates of standard errors and overstated estimated significance levels. We also report estimates using heteroskedasticity-robust standard errors but adopt another widely popular approach in panel

---

12. We used an F test to determine whether, as a group, the year dummies affected cross-temporal changes in homicide rates. Not surprisingly, the F statistic shows the year dummies are jointly statistically significant at the .01 significance level. Interestingly, Dezhbakhsh and Shepherd (2006) chose not to include year fixed effects in their homicide specifications, although their importance in minimizing omitted variable bias is well documented in the crime literature using the panel data approach. In fact, we are not aware of any published panel studies of crime that have tested (using an F test) and not found the year dummies to be highly significant as a group. Instead, the authors opted for the use of decade fixed-effects that control for the average homicide rate in each decade. The problem with this approach, however, is that it fails to take into account the fact that trends in homicide rates have varied largely within each decade since the 1960s (see also Donohue and Wolfers, 2005: 805–806). Not surprisingly, when Donohue and Wolfers (2005) reestimated Dezhbakhsh and Shepherd's specification examining the impact of state DP moratoriums (imposed by the 1972 Supreme Court decision in Furman v. Georgia) on homicide rates (see their Table 8, Column 2) while controlling for year fixed effects, the coefficient on the DP moratorium dummy variable was cut almost in half and no longer statistically significant.

studies, which is to enter a 1-year lag of the dependent variable in the specification to correct for serial correlation and to mitigate omitted variable bias (Beck and Katz, 1995, 1996). As is well known, per-capita variation in crime rates is greater in low-population states. In this case, the OLS estimator is no longer efficient. In an attempt to gain efficiency in our parameter estimates, we decided to use weighted least squares (WLS) regression, where the weights are the number of people who live in each state. Although population size may not correspond to the inverse of the error variances, the WLS estimator is likely to be more efficient than OLS. Importantly, even if the weights are not optimal or heteroskedasticity remains, the use of robust standard errors (clustered or not) will still provide for robust inference.

T A B L E   2

Summary Statistics

	Mean	Overall SD	Within–State SD
Dependent variables			
Homicides per 100K	6.290	3.572	1.680
Death penalty variables			
Death penalty dummy variable	0.719	0.450	0.172
Executions(t)	.703	2.803	2.096
Executions(t) per 1,000 prisoners(t)	.030	.094	.085
Executions(t)/death sentences (t–1)	.092	.320	.283
Executions(t)/death sentences(t–6)	.102	.348	.311
Executions(t) per 100K(t)	.012	.038	.033
Executions(t)/homicides(t–1)	.002	.008	.007
Policy control variable			
Right-to-carry concealed law dummy variable	.230	.416	.338
3X law dummy variable	.193	.391	.335
Sociodemographic control variables			
Unemployment rate	5.878	2.004	1.659
Employment rate	55,631.03	6,026.39	4333.71
Poverty rate	12.841	3.937	2.020
Real per-capita income	4.777	.920	.631
Percent persons ages 15 to 24	15.660	2.131	1.986
Percent persons ages 25 to 34	15.565	1.987	1.711
Percent persons ages 35 to 44	14.275	2.053	1.886
Beer shipments (31-gallon barrels) per 100K	75,303.87	13,108.02	5,442.83
Percent persons with college degree	20.954	5.497	4.165
Percent persons residing in metropolitan areas	65.887	21.322	3.249
Crack index, 1980 to 2000	0.948	1.264	1.075
Deterrence variables			
Prison deaths per 1,000 prisoners	128.504	155.616	72.785
Prisoners per 100K	257.066	149.761	113.874
Police officers per 100K	265.764	50.188	24.278

Notes. Descriptive statistics are for the 1977 to 2006 time period except where noted in text. Means and standard deviations are unweighted. The data sources are described in the text.

Next, we examined the stationarity of the homicide rate using the panel unit root test advocated by Im, Pesaran, and Shin (2003), hereafter indicated by IPS. Assessing the stationarity of homicide is important, as standard significance tests assume variables have the property that the mean and variance are constant over time (Moody, 2005). The IPS test is based on the null hypothesis that all homicide series are generated by unit-root processes. The alternative is that at least one homicide series is stationary. For panel unit-root tests, the lag length has to be chosen. The data are annual, so a lag length of 1 was chosen, but similar results were obtained when using lag lengths of 2 and 3 years. The results of the IPS test reject the null hypothesis of a unit root in homicide rates at the 1% level. The standardized t-bar statistics for the homicide rate are -4.21 ($p < .000$), -2.814 ($p < .002$), and -2.975 ($p < .001$) when the lag lengths are set equal to 1, 2, and 3 years, respectively. These results both replicate the mean-reverting property of the homicide rate reported by Moody (2005) and show that his results do not change with extended data. Because homicide seems to be a stationary process, there is no need to first-difference the variables (i.e., we estimate the regressions using levels of variables).

Finally, an examination of collinearity diagnostics reveals no serious collinearity problems for any of the DP measures, although it does affect some of the other explanatory variables. This occurs mainly for the socioeconomic variables that change slowly over time and are highly correlated with the state fixed effects. Thus, it is necessary to use caution in the interpretation of results for these slow-moving variables.

Table 2 lists the variables used in the regression models. In addition to the variable name and a brief description, the mean, and overall and within-state standard deviations are shown. Estimation was carried out in Stata, version 9.0 (StataCorp., College Station, TX).

Empirical Results

Table 3 presents the results of eight separate homicide estimations using regression procedures for panel data discussed in the previous section. The most notable features include using state and year fixed effects and linear state-specific time trends, and using WLS, where the weight is the state's share of the U.S. population. Heteroskedasticity-autocorrelation robust standard errors, which are clustered at the state level, are presented in parentheses under the regression coefficients. Each homicide model includes, one at a time, a measure that captures the presence of a DP statute, the frequency of executions, or the probability of execution for homicide. Because of space limitations, the regression coefficients for the state and year fixed effects and linear state-specific time trends are not shown.

Presence of the Death Penalty and Homicide

We begin the analysis by examining whether a baseline deterrent effect of the DP on homicide rates is attributable to the presence of an active DP statute. Because our study sample begins in the post-moratorium era, we cannot assess the effects on homicide rates of state DP statutes reenacted after the *Furman v. Georgia* (1972) decision (which became effective after the U.S. Supreme Court's decision on July 2, 1976) or those state statutes enacted during the year 1977

(e.g., CA, NC, and WY). Any deterrent impact caused by the presence of these laws is captured by the state fixed-effects variables. Nevertheless, we still can assess the effects of 11 separate changes to state DP statutes that have taken place from 1977 to 2006. Of these 11 policy changes, 8 were based on states enacting (in some cases, reenacting) a DP law (KS, 1994; MA, 1982; NH, 1991; NJ, 1982; NM, 1979; NY, 1995; OR, 1978; SD, 1979), 1 abolishing its DP statute (RI, 1984), and 2 suspending executions because of a Governor-ordered moratorium (IL, 2000; MD, 2002).[13] Similar to Donohue and Wolfers (2006), we use a binary dummy variable set equal to 1 when a state has an active DP statute and 0 otherwise.

TABLE 3

The Impact of the Death Penalty on Homicide Rates: Estimates from State Panel Data, 1977 to 2006

Variable	Dependent Variable: Annual Homicides per 100,000 State Population							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Death penalty dummy variable (t)	−0.416 (.507)							
Executions (t)	—	−0.007 (.014)	—	—	—	—	—	—
Executions (t−1)	—	—	0.006 (.008)	—	—	—	—	—
Executions (t) per 1,000 prisoners (t)	—	—	—	−0.102 (.593)	—	—	—	—
Executions (t)/death sentences (t − 1)	—	—	—	—	0.094 (.087)	—	—	—
Executions (t)/death sentences (t − 6), (t − 1)	—	—	—	—	—	−0.071 (.129)	—	—
Executions (t) per 100K (t)	—	—	—	—	—	—	−0.166 (1.399)	—
Executions (t)/homicides (t − 1)	—	—	—	—	—	—	—	1.799 (5.763)
Prison death rate	0.001 (.001)	0.002 (.001)	0.002 (.001)	0.002 (.001)	0.001 (.001)	0.0002 (.002)	0.002 (.001)	0.002 (.001)

13. Massachusetts abolished its DP in 1984, whereas the New York State Supreme Court ruled the state's DP law unconstitutional in June 2004. Rather than estimate separate effects for both the enactment and the abolition of these state laws, we follow the strategy of Dezhbakhsh and Shepherd (2006) and create a single law dummy variable that accounts for both policy changes simultaneously. For example, the DP law dummy variable for New York is coded 0 for the years 1977 to 1994, 1 for the years 1995 to 2003, and 0 for the years 2004 to 2006.

Shall issue law	−0.345	−0.317	−0.332	−0.324	−0.338	−0.413	−0.323	−0.328
	(.260)	(.263)	(.259)	(.258)	(.268)	(.270)	(.259)	(.259)
3X law	**1.044**	<u>1.085</u>	<u>1.101</u>	<u>1.093</u>	<u>1.097</u>	1.038	<u>1.092</u>	<u>1.097</u>
	(.500)	(.572)	(.569)	(.570)	(.567)	(.620)	(.571)	(.57)
Unemployment rate	**−0.124**	<u>−0.112</u>	<u>−0.116</u>	<u>−0.113</u>	−0.104	−0.041	<u>−0.114</u>	<u>−0.115</u>
	(.061)	(.062)	(.063)	(.060)	(.069)	(.099)	(.061)	(.062)
Employment rate × 100	0.003	0.004	0.004	0.004	0.002	−0.004	0.004	0.004
	(.007)	(.007)	(.007)	(.007)	(.008)	(.010)	(.007)	(.007)
Poverty rate	0.040	0.040	0.040	0.040	0.033	0.010	0.040	0.040
	(.040)	(.040)	(.041)	(.040)	(.041)	(.042)	(.040)	(.040)
Per-capita income	0.168	0.160	0.178	0.163	0.236	0.004	0.165	0.177
	(.718)	(.716)	(.718)	(.721)	(.734)	(.764)	(.716)	(.717)
Percent aged 15 to 24	0.118	0.105	0.105	0.105	0.221	<u>0.612</u>	0.105	0.105
	(.196)	(.190)	(.191)	(.191)	(.180)	(.215)	(.191)	(.190)
Percent aged 25 to 34	0.397	0.366	0.369	0.368	<u>0.423</u>	**0.812**	0.368	0.367
	(.255)	(.249)	(.248)	(.248)	(.239)	(.310)	(.248)	(.248)
Percent aged 35 to 44	−0.105	−0.111	−0.114	−0.113	−0.062	−0.026	−0.113	−0.114
	(.236)	(.231)	(.233)	(.231)	(.246)	(.329)	(.232)	(.233)
Prisoners per 100K, $(t − 1)$	**<u>−0.009</u>**	**<u>−0.009</u>**	**<u>−0.009</u>**	**<u>−0.009</u>**	**<u>−0.008</u>**	**<u>−0.009</u>**	**<u>−0.009</u>**	**<u>−0.009</u>**
	(.002)	(.002)	(.002)	(.002)	(.002)	(.002)	(.002)	(.002)
Police officers per 100K, $(t − 1)$	**−0.011**	**−0.012**	**−0.012**	**−0.012**	**−0.013**	**−0.017**	**−0.012**	**−0.012**
	(.005)	(.006)	(.006)	(.006)	(.006)	(.007)	(.006)	(.006)
Beer consumption × 100	0.004	0.004	0.004	0.004	0.002	<u>0.008</u>	0.004	0.004
	(.003)	(.003)	(.003)	(.003)	(.003)	(.004)	(.003)	(.003)
Percent college degree	0.254	0.251	0.257	0.254	0.174	0.290	0.253	0.255
	(.170)	(.175)	(.175)	(.174)	(.172)	(.293)	(.174)	(.174)
Percent metropolitan	−0.023	−0.021	−0.022	−0.021	−0.013	−0.051	−0.022	−0.022
	(.043)	(.042)	(.042)	(.042)	(.048)	(.090)	(.042)	(.042)
Sample size	1,499	1,499	1,499	1,499	1,449	1,149	1,499	1,499
Adjusted R^2	.94	.94	.94	.94	.94	.93	.94	.94

Notes. The dependent variable is the annual number of homicides per 100,000 state population. The study period is 1977 to 2006 except as noted in the text. Prison death data for Alaska in 1994 are missing. The regressions are weighted by state population. Although not shown, state fixed effects, year fixed effects, and state-specific linear trends are included in all estimations and are always significant as a group using an F test. Heteroskedasticity and autocorrelation robust standard errors are reported. Coefficients that are significant at the 10% level are <u>underlined</u>. Coefficients that are significant at the 5% level are in **bold**. Coefficients that are significant at the 1% level are in **<u>bold and underlined</u>**.

The results for the DP law dummy variable are presented in column 1. Contrary to the findings reported by Dezhbakhsh and Shepherd (2006) and Mocan and Gittings (2003), but consistent with those reported by Donohue and Wolfers (2005), our results indicate no relationship between the activity status of the DP and homicide. Although the coefficient on the DP law dummy variable is in the negative direction, which is consistent with the DP deterrence hypothesis, it is not significantly different from 0. Next, we reestimated the specification in column 1 but lagged the DP law dummy 1 year to account for potential delays in the diffusion of information about a DP regime and, thus, its ability to alter prospective murderers' awareness of the possibility of being executed for murder. Lagging the DP law dummy variable by

1 year also helps to mitigate potential simultaneity bias if increases in the homicide rate lead state policy makers to adopt DP legislation. Although not shown, the coefficient on the DP law dummy was even smaller in absolute value than the current-year value and remained statistically insignificant. In all, the results of the dummy variable analysis provide little systematic evidence that the mere possibility of being executed for murder serves as an effective deterrent to potential murderers, at least not during the post-moratorium era.

Admittedly, one drawback to this analysis is that most states with currently active DP statutes (re)enacted them prior to 1977. As a result, we only could assess the potential deterrent impact of the presence of the DP on homicide in a minority of DP states. Perhaps more importantly, the dummy variable approach cannot address what most DP deterrence scholars consider to be the more relevant empirical question related to the DP deterrence hypothesis: Do higher levels of execution risk produce stronger deterrent effects? As discussed, most DP scholars argue that the strongest deterrent effects of the DP are likely to be linked to its application in practice. For example, if the cognitive link in a potential murderer's mind is the actual risk of execution for homicide, rather than the possibility of execution, then the appropriate independent variable in a DP study is the frequency of executions or the probability of execution for homicide rather than its presence. Indeed, most of the DP papers reviewed in Table 1 find the strongest support for the DP-deters-homicide hypothesis when examining the link among the frequency of executions (e.g., Dezhbakhsh and Shepherd, 2006), probability of execution (e.g., Mocan and Gittings, 2003), and homicide rates rather than the legal status of the law. As a result, we turn our attention on the relationship between the application of the DP and homicide using measures most commonly employed in recent DP deterrence research.

Number of Executions, Probability of Execution, and Homicide

Columns 2 through 8 report the results of seven different estimations using the same exact model specification employed in Column 1 but replace the DP law dummy variable with either the frequency of executions or the probability of execution for a given cohort of incarcerated murderers. Again, we emphasize that what is important from a deterrence perspective is a prospective murderer's *perceived* risk of execution for homicide. Obviously, direct measures of perceived risk of execution at the aggregate level are nonexistent, and thus, DP scholars have been forced to use aggregate-level measures of actual execution risk as proxies for the aggregate perceived risks. To the extent that aggregate-level measures of actual execution risk have a significant positive association with the perceived risks, the objective risks should provide a satisfactory proxy for the perceived risks (but see Kleck et al., 2005).

The results in Column 2 are based on Dezhbakhsh and Shepherd's (2006) preferred measure of simply using the total number of executions in a state–year. Contrary to the findings reported by Dezhbakhsh and Shepherd (2006), our results indicate that increasing the scale of executions does not lead to greater deterrent effects by "sending a message" to potential murderers of the state's willingness to execute persons convicted of homicide. The sign on the execution variable

is negative but far from significant. Importantly, the coefficient on the execution measure is roughly one twentieth the size reported by Dezhbakhsh and Shepherd (2006; see their Table 8, Column 1). Lagging the execution variable 1 year to account for potential delays in the transmission of this deterrence message produced substantively similar results (Column 3). Needless to say, the vastly different results obtained by Dezhbakhsh and Shepherd (2006) and the current study for the execution measure was a source of concern for us. As one of our many robustness checks, we altered our baseline specification to resemble more closely the estimation method used by the authors to identify the source of the differences. We believe the results obtained by Dezhbakhsh and Shepherd (2006) for the execution measure were a by-product of omitted variable bias and failing to adjust standard errors for the presence of serial correlation.

Column 4 reports estimates using executions carried out in year t per 1,000 state prisoners as recommended by Katz et al. (2003). Interestingly, the sign and value of the coefficient for the execution variable are identical to those reported by Katz et al. (2003; see their Table 2, Column 6), although we report standard errors much larger than theirs. Regardless, our finding of no significant relationship between the risk of execution and the homicide rate is consistent with that reported by Katz et al. (2003). Columns 5 and 6 report estimates using slightly different variants of ratio variables capturing the actual objective probability of execution for homicide. The first ratio measure is similar to the one used by Zimmerman (2004) and is the number of executions carried out in year t divided by the number of persons sentenced to death row in year $t - 1$. The second ratio measure, which was employed by Mocan and Gittings (2003), is similar to the first measure except that the denominator is death sentences in year $t - 6$.[14] The theoretical justification for lagging the denominator by 1 and 6 years was discussed above. In both cases, the coefficients on the execution risk measures are far from significant and in the case of the former measure it is actually in the unexpected positive direction.[15] We also created risk of execution measures assuming the average wait on death row was 4 and 5 years. Although not shown, the results for these alternative versions were largely similar to those obtained when deflating the denominator by 6 years.

The last two execution measures in Table 3 were employed exclusively by Donohue and Wolfers (2005). Column 7 reports estimates using executions in year t per 100,000 population, whereas the results in Column 8 are based on estimations using executions divided by homicide in year $t - 1$.[16] Once again, these execution risk measures fail to reveal any significant negative relationship between the risk of execution and homicide. The coefficients for both execution

14. Following Mocan and Gittings (2003), we lag the risk of execution measure by 1 year to be more consistent with their specification.

15. Donohue and Wolfers (2005) employ a slightly different variant of Zimmerman's execution measure in which they lag the numerator by 1 year to mitigate simultaneity bias. We also tried this measure, and it produced results largely similar to those obtained using the Zimmerman measure.

16. Donohue and Wolfers (2005) use executions per 100,000 population because of the scaling problems discussed earlier when using only the sheer number of executions.

risk measures are far from significant, and the sign on the latter measure is in the unexpected positive direction.

Other Notable Findings

Although it is not the focus of the current study, the performance of the specific control variables in Table 3 are worth a brief mention because many of them are considered important correlates of homicide, and the results for these explanatory variables may, at least for some readers, speak volumes with regard to the reliability of the findings for the DP deterrence measures. First, we find no evidence that increases in the presence of young adults is associated with higher rates of homicide. These results support recent empirical works by Levitt (1999) and Marvell and Moody (2001). With respect to the policy variables, the adoption of 3X laws is positively correlated with higher homicide rates. This finding is consistent with Kovandzic et al. (2002) and Marvell and Moody (2001). Similar to Ayres and Donohue (2003) and Kovandzic, Marvell, and Vieraitis (2005), we find no evidence to support a deterrent effect of the passage of shall-issue concealed handgun laws. We also find no evidence that worsening prison conditions, as proxied by prison deaths, reduces the homicide rate. These results parallel those reported by Katz et al. (2003), although the authors reported statistically significant decreases in almost all cases for the violent crime rate and in a few cases for the property crime rate. Last, police levels and prison population growth are both significantly related to lower homicide rates. This finding largely mirrors those reported in other state panel studies (e.g., Katz et al., 2003; Levitt, 1996; Marvell and Moody, 1994, 1996; Zimmerman, 2004).

Does a Two-Way Relationship Exist Between Execution Risk and Homicide?

Recent DP deterrence authors (e.g., Mocan and Gittings, 2003; Zimmerman, 2004) have suggested that coefficients for execution risk measures might suffer from simultaneity bias if increases in homicide rates heighten public fear of crime and, in turn, encourage prosecutors to seek the DP more often and make judges less likely to overturn death sentences imposed by juries. Research also suggests that public opinion and elections influence judicial decision making, with sentencing becoming more punitive as elections near (Brace and Boyea, 2008; Huber and Gordon, 2004). If this is the case, and such contemporaneous homicide effects are ignored, then simultaneity bias would cause OLS estimates to underestimate the deterrent effect of execution risk on homicide. Mocan and Gittings do not formerly address the simultaneity problem; rather, they attempt to mitigate the problem by lagging their execution risk measure by 1 year. The justification for this approach is that contemporaneous homicide rates cannot influence the execution risk for the prior year. In practice, however, execution risk measures are unlikely to suffer from this form of endogeneity bias because of the lengthy time lag between the offense date and the execution date. For example, 2 of the 1,004 (or 0.2%) offenders sentenced to death row between 1977 and 2005 were executed in the same year they were sentenced (senior author's analysis of Crime in the United States data set).

Zimmerman (2004: 173) presented another argument for potential simultaneity between execution risk and homicide, which he refers to as the "lethality effect" of the DP. He suggested that some "rational offenders" might decide to eliminate potential victims and witnesses if doing so reduces their risk of execution. Zimmerman (2004) explained the potential consequences of the "lethality effect" of capital punishment:

> If such a lethality effect of capital punishment is operative, estimates of the deterrent effect of capital punishment will be biased upwards since reverse causation operates in the negative direction. Correcting for simultaneity in this case would result in a smaller estimated deterrent effect.

Although the lengthy lag between the offense date and execution make the lethality effect argument tenuous at best, Zimmerman (2004) was correct in pointing out that OLS estimates for the probability of execution risk will suffer from simultaneity bias if lethality effects that take place in a given year concomitantly lead to lower levels of execution risk. The reason is that the regressor, execution risk, is itself endogenous in a system of simultaneous equations, which makes it correlated with the error term in the homicide model. As Zimmerman noted, the coefficient for the execution risk variable will be biased negatively because the killing of witnesses lowers the probability that some offenders will be arrested, convicted, and subsequently executed. What we find puzzling, however, is that Zimmerman attempted to correct for this potential simultaneity problem using instrumental variable (IV) methods when by his own accounting the small, nonsignificant OLS results he reports for the execution risk variables (current year or lagged 1 year) were already biased in favor of support of the DP deterrence hypothesis. In this case, using IV methods to purge the homicide equation of simultaneity bias based on the lethality effects would only have served the purpose of making the IV estimates for execution risk less negative than the OLS estimates, which were already close to 0. Importantly, however, this is not what Zimmerman (2004) found when implementing IV methods. Instead, Zimmerman reported IV estimates for execution risk that are roughly 15 times larger in the negative direction than the OLS estimates.[17] Such a result is consistent with severe reverse causality operating in the positive direction; this finding completely contradicts Zimmerman's (2004) lethality effect argument. The most likely explanation for the large divergence between the OLS and IV estimates for execution risk is that the instrumental variables used by Zimmerman (2004) to instrument for execution risk were invalid (i.e., negatively correlated with

---

17. Zimmerman (2004) examined execution risk with a pair of dummy variables. The first dummy variable denotes the presence of a botched execution in the previous year, whereas the second dummy variable denotes the removal of an inmate from death row in the previous year. The author suggested that both of these events should lead to fewer executions in the subsequent year but have no direct impact on homicide rates in the next year. He also treated the probability of arrest and receiving a death row sentence after conviction as endogenous regressors in his IV estimations and evaluates these two deterrence variables using three additional instruments.

the error term).[18] In all, the evidence from IV estimates, in our opinion, offers no support for the DP deterrence hypothesis.

David Greenberg and Gary Kleck have brought to our attention perhaps the most plausible mechanism by which homicide rates might reverse cause execution risk in the same year: It might be riskier politically for governors or parole boards to commute sentences in years with a greater number of homicides, and conversely, in years with fewer murders, it might be easier for such parties to show their sense of compassion by commuting near-term executions. If either situation occurred in practice, then a simultaneous relationship would exist between the homicide rate and execution risk. To examine this possibility, we computed the within-state bivariate correlation (i.e., we controlled for state fixed effects) between homicide rates and the total number of inmates on death row who had their death sentences commuted from 1977 to 2005. Although the Pearson correlation coefficient was in the expected negative direction ($r = -.010$), it was small and far from being significantly different from 0 ($p = .702$). Similar results were obtained when we used the total number of homicides instead of the homicide rate ($r = -.021/p = .394$). These results, coupled with the facts presented above regarding the lengthy lag between the offense date and the date of execution, suggest current-year execution risk is an exogenous event that has little or nothing to do with current-year homicide rates.

Testing the Sensitivity of the Results

As Beck and Katz (1996) noted, there is no magic bullet estimator for panel data, and analysts who use such data must make many difficult decisions throughout the statistical modeling process. They suggested that modeling decisions should be based on both relevant theory and the methodological literature on panel data. We agree with Beck and Katz (1996) and consider the statistical fixes selected here to be the preferred "cures" for the problems present in our panel data set. In addition, we believe the DP measures used here most closely represent the plausible theoretical processes by which the DP is supposed to deter homicide. However, we also realize

---

18. Zimmerman (2004) reported the results of several tests to demonstrate to readers that all three deterrence variables, including execution risk, are in fact endogenous regressors, and to establish the relevance and validity of the instruments used for the deterrence variables. Unfortunately, the tests suffer from technical problems or are seriously flawed. For example, he used outdated Sargan and Durbin-Wu-Hausman versions of overidentification and exogeneity tests to establish instrument validity and the endogeneity of the deterrence regressors, respectively. The problem with these tests, however, is they are invalid in the presence of heteroskedasticity and nonindependence of error terms even though both forms of errors are almost always encountered by researchers using state panel data. Indeed, Zimmerman reported the presence of heteroskedasticity in his data. Furthermore, his suggestion that the excluded instruments were relevant (i.e., correlated with the endogenous deterrence regressors) based on the statistical significance of the first-stage F statistics for the instruments as a group is incorrect. A consensus has been reached in the econometric literature that it is not enough for the F statistic to be significant at conventional levels; higher values are required. Staiger and Stock (1997) recommended an F statistic of at least 10 as a "rule of thumb" for the IV estimator. In any event, one cannot determine whether a model is (under)identified using an F test when there are multiple endogenous regressors, as this requires estimation of the rank of the covariance matrix of regressors and instruments (Kovandzic, Schaffer, and Kleck, 2005). Two statistics that have been suggested for this purpose are the Cragg-Donald statistic and Anderson's canonical correlations statistic. Neither of these tests was reported by Zimmerman.

T A B L E   4

The Impact of the Death Penalty on Homicide Rates: Alternative Statistical Models

Model Specification	Death Penalty Dummy Variable (t)	Executions (t)	Executions (t) per 1,000 Prisoners (t)	Executions (t)/ Death Sentence (t – 1)	Executions (t)/ Death Sentence (t – 6), (t – 1)	Executions per 100,000 Population	Executions(t)/ Homicides (t – 1)
			Dependent Variable: Annual Homicides per 100,000 State Population				
1. Estimates from Table 3	−0.416	−0.007	−0.102	0.094	−0.071	−0.166	1.799
	(.507)	(.014)	(.593)	(.087)	(.129)	(1.399)	(5.763)
2. Time period, 1977 to 2000	−1.32	−0.008	−0.388	0.003	−0.154	−0.664	2.409
	(.846)	(.019)	(.726)	(.175)	(.144)	(2.055)	(10.212)
3. Drop irrelevant controls	−0.423	−0.007	−0.151	0.082	−0.072	−0.268	1.198
	(0.509)	(0.014)	(0.629)	(0.086)	(0.142)	(1.470)	(5.950)
4. Allocate executions using Mocan & Gitting's (2003) algorithm	—	−0.003	−0.072	0.249	−0.076	0.196	4.29
	—	(.019)	(.799)	(.119)	(.193)	(1.570)	(6.234)
5. Log-level model	−0.041	0.001	0.034	0.018	0.001	0.105	0.561
	(.048)	(.002)	(.064)	(.012)	(.012)	(.152)	(.765)
6. Unweighted by population	0.009	−0.025	−0.603	−0.050	−0.202	−1.710	−7.156
	(.451)	(.017)	(.434)	(.096)	(.122)	(1.242)	(3.768)
7. Robust standards errors, no cluster adjustment	−0.416	−0.007	−0.102	0.094	−0.071	−0.166	1.799
	(.243)	(.009)	(.446)	(.089)	(.108)	(.993)	(4.147)
8. Enter lagged DV, dynamic panel model	−0.161	−0.006	−0.196	−0.016	−0.048	−0.338	4.324
	(.171)	(.008)	(.367)	(.049)	(.095)	(.844)	(3.585)
9. PCSEs, lagged DV	−0.161	−0.006	−0.196	−0.016	−0.048	−0.338	4.324
	(.179)	(.012)	(.357)	(.077)	(.091)	(.889)	(4.326)
10. Control for crack epidemic	−1.167	−0.009	−0.433	−0.028	−0.167	−0.916	1.334
	(.836)	(.019)	(.680)	(.165)	(.149)	(1.941)	(9.989)

11. Drop year dummies	-0.752	**-0.045**	-1.543	-0.072	-0.116	**-3.808**	-18.244
	(.737)	(.009)	(.814)	(.104)	(.088)	(2.077)	(11.19)
12. Drop state-specific trends	-0.562	**-0.058**	-0.298	-0.152	-0.053	-1.335	-5.275
	(.352)	(.020)	(.991)	(.176)	(.240)	(3.072)	(16.514)
13. Drop year fixed effects and state–specific trends	-0.558	**-0.076**	-1.670	-0.303	-0.199	-3.951	-19.503
	(.562)	(.022)	(1.395)	(.264)	(.280)	(3.912)	(21.524)
14. Add decade fixed effects, drop year dummies, and state– specific trends	-0.555	**-0.062**	-1.284	-0.211	-0.160	-3.033	-12.754
	(.550)	(.022)	(1.256)	(.233)	(.276)	(3.644)	(19.351)

Notes. The data and specifications are identical to those used in Table 3 except as noted. Coefficients that are significant at the 10% level are underlined. Coefficients that are significant at the 5% level are in **bold**. Coefficients that are significant at the 1% level are in **bold and underlined.**

that some readers may consider our research agenda driven and believe, albeit incorrectly, that we have kept beneficial effects of the DP on homicide rates hidden from readers. As a result, we subjected the specifications in Table 3 to a battery of robustness checks using reasonable specification changes, such as (1) limiting the study period through 2000 to be more consistent with recent state panel DP papers, (2) dropping irrelevant control variables to avoid inflated standard errors for execution risk measures, (3) implementing the algorithm developed by Mocan and Gittings (2003) to prorate executions based on the month in which they occurred, (4) using a log-level model to determine whether this functional form provides a better fit to the data, (5) using alternative cures for problematic error variances in panel data, and (6) accounting for other sources of potential omitted variable bias (e.g., crack epidemic). We also report results for the risk of execution measures using model specifications that more closely resemble those employed in recent DP papers in an attempt to reconcile the conflicting evidence for several DP measures. Coefficient estimates for each of the seven execution risk measures are presented in separate columns in Table 4, with each row representing an alternative specification. For reference, the results for the execution risk measures reported in Table 3 where the dependent variable is annual homicides per 100,000 state population are reproduced.

Because the current study extends the analysis to include additional years (beyond 2000) not covered in recent state panel DP papers, several referees and the editor suggested we establish how important these additional 6 years are to the results. Row 2 of Table 4 presents results that limit the time series through 2000. The estimates of the effects of the DP on the homicide rate for the 1977 to 2000 sample are generally larger in the negative direction than those obtained during the entire sample period (1977 to 2006), but most are still much smaller in magnitude when compared directly with those reported in recent DP papers (using either pre- and post-Gregg data or only the latter) where similar measures were employed. More importantly, even when the models are estimated with the shorter time period, none of the DP measures are associated with lower rates of homicide, at least not at conventional significance levels. It seems, then, that lengthening the time series by 6 years is not responsible, at least not entirely, for the large differences obtained between the current study and the cadre of recent economic studies that report robust deterrent effects.

An anonymous reviewer suggested that the null results for the DP measures may be a by-product of including too many theoretically relevant but empirically irrelevant control variables in the baseline regression models. In other words, it is possible that the regression models were overparameterized because of the inclusion of too many regressors not significantly related to cross-temporal changes in homicide. Although the inclusion of irrelevant regressors is generally considered a "small" statistical problem (i.e., it does not lead to biased OLS estimates for the DP variables), it can lead to inflated estimated standard errors (i.e., inefficiency) if the irrelevant variables are correlated with measures of execution risk. To examine this possibility, we followed the referee's recommendation of implementing a procedure suggested by Rao (1971). The procedure entails dropping control variables from the regression model with t-ratios less than

1 and then verifying they are not jointly significant with a standard *F*-test. In our case, this led to the dropping of the following six variables from the regression model, as they consistently produced *t*-ratios of 1 or less across the various model specifications: employment rate, poverty rate, per-capita income, percent ages 15 to 24, percent ages 35 to 44, and percent metropolitan. *F*-tests also indicated these six variables were not jointly significant as a group and could be safely dropped from the model. Parameter estimates for the DP measures without these six control variables entered into the regression model are reported in Row 2 of Table 4. As viewed in Row 2, the results for the DP measures are largely similar to those reported in Table 3. It seems, then, that the dropped variables, although not significantly related to homicide rates, were also not significantly related to the presence of the DP or execution risk. In sum, we find no evidence that the nonsignificant results for the DP measures were a consequence of "overburdening" the baseline regression model with irrelevant regressors.

Next, we examined whether the null results for the execution risk measures were a by-product of our failing to take into account the timing of the execution event when calculating the risk measures. Similar to most DP studies, our measurement method assumes that executions taking place at any time of the year can influence the number of homicides in the same year. Mocan and Gittings (2003) argue this is not a reasonable assumption to make, as nearly half of all executions in the post-moratorium era have taken place between the months of July and December. As they correctly point out, an execution taking place in December of a given year cannot reasonably be expected to influence the homicide rate for that same year. Moreover, the authors maintain it is important from a theoretical perspective to create execution risk measures that best approximate the economic model of crime. Because economic theory indicates the timing of an execution event should matter to a prospective murderer when calculating their risk of execution, the most theoretically relevant measure of execution risk should take into account the month in which the execution took place. To examine this possibility, we altered the numerator of the execution risk measures using an algorithm developed by Mocan and Gittings (2003). The algorithm prorates executions based on the month in which they occurred.[19] For example, executions that take place in April and November of a given year count as 9/12 and 2/12 of an execution for the current year and 3/12 and 10/12 of an execution for the subsequent year, respectively. Row 3 of Table 4 reports the results of estimations when applying Mocan and Gitting's algorithm to compute the risk of execution measures. Most DP measures maintain their sign and remain insignificant. The only exception occurs for the ratio of executions to death sentences meted out in the previous year, which is now significant at the .05 level but in the unexpected positive direction. The null results reported here, therefore, do not seem to be an artifact of failing to account for the timing of execution events.

The specifications reported in Table 3 assumed a linear relationship between the risk of execution and homicide. Row 4 of Table 4 reports the results using a log-level model in which homicide is expressed in natural logarithms. This functional form examines whether an increase

---

19.   We are indebted to Mark Schaffer for showing us how to implement this algorithm in Stata.

in the risk of execution will lead to larger reductions in homicide when the homicide rate is high. In all cases, qualitatively similar results were obtained for the DP measures.

As another robustness check, we reestimated the specifications in Table 3 without weighting the regressions by state population. Because smaller states have larger error variances (for reasons noted), the resultant estimates for explanatory variables can no longer be considered efficient (i.e., precise), as all data observations are weighted equally. Nevertheless, estimating the regression models without weighting is a useful robustness check. The results of the unweighted regressions are presented in Row 5. With the exception of the ratio of executions to lagged homicides measures, which is statistically significant but only at the .10 level, the coefficient estimates for the DP measures remained statistically insignificant.

The next set of results in Table 4 examines the sensitivity of the standard errors when failing to correct for serial correlation or attending to the issue in alternative ways. In Row 6, we report heteroskedasticity-robust standard errors. As explained, these standard errors are robust to general forms of heteroskedasticity but not serial correlation. As evidenced by Donohue and Wolfers (2006), failing to correct standard errors for serially correlated errors can lead to much different conclusions for the DP measures.[20] Take, for example, the conclusion we would have necessarily drawn for the DP law dummy had we relied on these standard errors. As viewed in row 6 of Table 4, the coefficient for the DP law dummy is statistically significant but only because the standard errors are grossly underestimated and only at the more generous .10 level. We concur with the growing sentiment in economics that standard errors in panel data analysis should, at minimum, be "fixed up" for general heteroskedasticity and serially correlated errors.

Of course, serial correlation issues can be addressed in other ways besides "fixing up" the standard errors. Row 7 reports heteroskedasticity-robust standard errors, but instead of using a clustering correction, we follow the strategy advocated by pioneers in panel data methodology, Nathaniel Beck and Jonathan Katz (1994), of entering a 1-year lag of the homicide rate in the specification to correct for serial correlation.[21] The benefit of this approach is that does not treat

---

20. Donohue and Wolfers (2006) demonstrated that the standard errors used by Dezhbakhsh and Shepherd were also severely underestimated, as the authors failed to adjust them for serial correlation. At the time of their writing, the authors speculated that Dezhbakhsh and Shepherd must have mistakenly used ordinary OLS standard errors instead of what they claimed to be "standard errors corrected for possible clustering effects—dependence within clusters (groups)." It turns out, however, that Dezhbakhsh and Shepherd did use cluster robust standard errors but opted to cluster by "year" instead of "states" (Dezhbakhsh and Rubin, 2007). Our reading of the Dezhbakhsh and Rubin paper suggested that the authors believed they were being criticized by Donohue and Wolfers (2006) for not adjusting standard errors to correct for spatial correlation problems, but this was not the case. Instead, Donohue and Wolfers (2006) were concerned that the authors had failed to adjust the standard errors for potential serial correlation in the data, which as it turns out, they had not done.

21. It is well known that panel estimates for explanatory variables are likely to be biased in the presence of fixed effects and lagged dependent variables, especially when t is in the single digits. Research by Beck and Katz (1996) suggests, however, that the OLS estimator actually performs nicely for longer panel data and should be preferred over alternative estimators (e.g., Anderson-Hsiao estimator) proposed for panel data with fixed effects and lagged dependent variables, especially when T is greater than 20, as is the case here.

serial correlation as a nuisance; rather, it allows the analyst to model the processes generating the serially correlated errors by taking into account the dynamic aspect of the data. Including lagged dependent variables in the model also aids in alleviating omitted variable bias by controlling for omitted lagged predictors of the homicide rate (Marvell and Moody, 1996). The specification reported in Row 8 also includes a lagged dependent variable to control for serial correlation but employs the use of "panel corrected standard errors" (PCSE) to tweak the standard errors for additional problems—group-wise heteroskedasticity and temporally correlated errors (Beck and Katz, 2004). As noted by Beck and Katz (2004), heteroskedasticity-robust standard errors do nothing to remedy these likely problems in panel data analysis. As viewed in Rows 7 and 8, these alternative approaches for dealing with problematic error structures produces coefficient estimates for the DP measures that are qualitatively similar to those reported in Table 3; that is, most are negative, but none are significantly different from 0.

One possible explanation for our null results is that we omitted from the homicide specifications a factor that is positively correlated with both execution risk and the homicide rate. If such a factor did exist, it may have suppressed, at least partially, the negative effects of capital punishment on homicide rates. One possible suppressor variable is the crack epidemic, which is considered by most criminologists to be the main culprit for the soaring homicide rate among adolescents and young adults from the mid-1980s to the early 1990s (e.g., Blumstein, 1995). If states experiencing the brunt of the crack epidemic and the concomitant rise in homicide were more likely to respond by adopting DP legislation or more aggressively pursuing death sentences, then this would have provided the set (i.e., positive impact on both risk of execution and homicide rates) of positive associations needed to suppress the deterrent effects of risk of execution on homicide. To examine this possibility, we entered the crack index variable into the homicide specifications as a control variable. Again, the crack index variable was only available through 2000, so this shortened the length of the time series by 6 years. As viewed in Row 9, it seems the crack epidemic is not a suppressor variable in the DP law–homicide relationship. Although the coefficients on the DP measures are mostly negative and generally larger than those reported in Table 3, the standard errors are also much larger in size and, thus, are statistically insignificant.

The last set of specifications reported in Table 5 show the results of estimations in which the year fixed effects, state-specific trends, or both are dropped as control variables. As one might expect, dropping either group of proxy variables (or both) from the specification generally produces results more favorable to the DP efficacy hypothesis, especially for Dezhbakhsh and Shepherd's (2006) preferred number of executions measure. As viewed in Rows 11 through 13, Column 2, the coefficient estimates for the frequency of execution measure are substantially larger than those reported in Table 3 and are now highly significant, although they remain roughly half the size of those reported by Dezhbakhsh and Shepherd. Through the detective work of Donohue and Wolfers (2006), we eventually came to learn that Dezhbakhsh and Shepherd (2006) did not include year fixed effects or state-specific trends in any of the 96 regression models presented,

although their importance in minimizing omitted variable bias is well documented in panel studies of crime (Marvell and Moody, 1996).[22] Instead, the authors used what they refer to as "decade-specific dummy variables" to control for the average homicide rate in each decade. The problem with this approach, however, is that it fails to take into account that trends in homicide have varied largely within each decade since the 1960s (see also Donohue and Wolfers, 2006: 805–806).[23] Indeed, when we reestimated a regression model while controlling for decade fixed effects but without controlling for year fixed effects or state-specific trends, the coefficient on the frequency of execution measure remained highly significant in the negative direction (see Row 14, Column 2). It seems, then, that decade fixed effects are clearly inferior to year fixed effects and state-specific trends when it comes to ameliorating omitted variable bias. We believe year fixed effects and state-specific trends should be included in the model specification when their estimated effects are large and significant, as is the case here. Perhaps more importantly, the results obtained for the frequency of execution measure when these proxy variables are not included in the regression model are almost assuredly spurious.

To summarize, the results in Table 4 confirm that the results for the DP measures are largely insensitive to changes in our choice of functional form, weighting schemes, procedures used for correcting problematic error variances, and when controlling for one of the most significant historical events in the post-moratorium era linked to dramatic increases in homicide rates—the crack epidemic.

Discussion

Our finding that the DP is not a significant deterrent to homicide is consistent with research by economists such as Katz et al. (2003) and Donohue and Wolfers (2005), but it differs sharply from the strong prodeterrence findings of many recent DP studies conducted by economists. Given that most of these studies, using sophisticated econometric methods, have found strong support for the DP deterrence thesis, and considering the history of DP research and its impact on crime policy, we carefully examine the implications of our findings.

The most likely explanation for the divergence between our largely null findings and studies reporting robust deterrent effects that result from increases in execution risk is the failure

---

22.   We are not aware of any published studies that have tested and not found the year fixed effects or state-specific trends to be highly significant as a group. As noted, when we tested the significance of the year fixed effects and state-specific trends as groups using the classic F-test, the null hypothesis of no effects was rejected for each set of proxy variables across all model specifications reported in Table 3.

23.   Donohue and Wolfers (2006) reported that the coefficient and standard error obtained by Dezhbakhsh and Shepherd (2006) for the execution measure remained largely unaltered when they replicated the authors' preferred specification while controlling for year fixed effects (see their Table 5, Column 2). Interestingly, the standard error for the execution measure increased dramatically when they dropped Texas from the sample and was no longer significantly different from 0. They do not report estimates using state-specific trends as control variables, but they do report that Dezhbakhsh and Shepherd's (2006) findings of higher homicide rates during state DP moratoriums (mostly imposed by the 1972 Supreme Court decision in Furman vs. Georgia) completely vanished when they entered year fixed effects into their model specification (see their Table 2, Column 3). Specifically, the coefficient on the DP moratorium dummy variable was cut almost in half and no longer statistically significant.

of the latter to (1) address adequately omitted variable bias by failing to include year dummies and or state-specific trends in the regression model, (2) adjust standard errors to correct for serial correlation, and (3) use reliable and valid instruments to address potential simultaneity bias between execution risk and homicide. Interestingly, the inclusion of additional variables to control for the effects of concurrent confounding policy factors (i.e., 3X and RTC laws) and historical events (i.e., crack epidemic), which have been empirically linked with cross-temporal changes in homicide rates, and extending the sample period included in recent DP studies by 6 years (from 2000 through 2006), proved to be of little consequence.

Our finding that the DP is not a deterrent to homicide is probably not surprising to most criminologists or others knowledgeable about the existing research on the death penalty, offender decision making, and/or the nature of homicide events. In contrast to the view of offender decision making modeled in recent DP research, which relies heavily on price theory and on the assumption that the process applies equally to potential offenders and the situational contexts in which they choose to commit homicide, criminological research on offenders suggests that the process is much more complex. Our null findings, therefore, are largely consistent with the considerable body of research on offender decision making and with research on the nature of homicide. As such, we turn our discussion to the criminological research that can help make sense of our findings and provide guidance for future deterrence research.

Nature of the Criminal Calculus

Although many recent studies have relied on the Beckerian model of criminal decision making, assuming that offenders are rational and thus likely deterred when the "costs" (e.g., the death penalty) are greater than the "benefits" (e.g., killing someone to settle a dispute), criminological theory and empirical research suggest that the process is much more complicated. We do not suggest, however, that no criminals follow the rational cost–benefit model proposed by Becker, but that most, including prospective homicide offenders, likely do not. Contemporary versions of rational choice theory put forth a more multifaceted view of offender decision making than early versions such as Becker's (1967). For example, in contrast to the image of a rational calculating criminal, Cornish and Clarke (1986) portray criminals' rationality as bounded or limited. As such, offenders do not always succeed in making the "best" decisions (i.e., forgoing criminal behavior) because they, like the rest of us, rarely have all the facts about the potential costs and benefits of an action. Often, as is most commonly the case with homicides, choices to commit crimes are made hastily or in the heat of the moment rather than after careful planning and deliberation. Moreover, many offenders make decisions under the influence of illegal drugs or alcohol (Bureau of Justice Statistics, 2006; National Institute of Justice, 2003) and are not lucid in their thoughts or behaviors at the time a crime occurs. Thus, offenders, especially violent offenders, are rarely cold and calculating but rather are entrenched in a lifestyle of drugs, alcohol, and desperation (Jacobs, 1999; Wright and Decker, 1997). Perhaps the most relevant question regarding the deterrent effects of legal sanctions is determining to what degree, if any,

offenders consider these punishments prior to engaging in illegal behaviors. Although they disagree on what specific information offenders use in their decision making, many recent DP studies assume that offenders actually do use information about legal sanctions and that these assumptions are accurate. It is on this point that criminological research also has had much to contribute, and it is toward this issue that we now turn.

A substantial body of criminological research exists that provides rich details about the motivations and causes of crime, the situational dynamics of the criminal event, and the nature of the criminal calculus (e.g., Bennett and Wright, 1984; Bourgois, 1995; Cromwell and Olson, 2004; Jacobs, 1999; Rengert and Wasilchick, 1985; Shover, 1996; Tunnell, 1992; Wright and Decker, 1994, 1997). This line of inquiry has led to a deeper understanding of the decision-making process of offenders, including whether and how criminals evaluate the risks of getting caught, convicted, and sentenced. Qualitative interviews with offenders have shown that they are more likely to focus on the rewards rather than on the risks of their actions. Moreover, even if they do consider the risks, they tend to focus on immediate versus long-term or worst-case risks, such as death. Recent ethnographic evidence supports the view that offenders' rationality is bounded and made within a social context. Decisions are often made in the context of a criminal lifestyle described as "then and there" (Wright and Decker, 1997) or an "unending party" (Shover, 1996). Decisions are rarely, if ever, made after careful reflection on and consideration of the potential costs and benefits.

Evidence from interviews with both active and captured offenders suggests they do not dwell on the potential consequences of their actions and thus rarely consider the possibility of arrest and imprisonment (Bennett and Wright, 1984; Copes and Vieraitis, 2008; Cromwell and Olson, 2004; Jacobs, 1999; Rengert and Wasilchick, 1985; Shover, 1996; Shover and Honaker, 1992; Tunnell, 1992). For instance, Cromwell and Olson's (2004) study of burglars found that consideration of long-term risk was almost nonexistent in their decision process. Based on his interviews with crack dealers, Jacobs (1999) found that the certainty of punishment, not severity, was foremost in sellers' minds but that many offenders believed that they would never get caught. As such, Jacobs (1999: 116) concluded that "the fear of harsh punishment is unlikely to work as a crime prevention mechanism in light of the techniques criminals use prospectively to evade apprehension" (see also Bourgois, 1995; Cherbonneau and Copes, 2006). Even when criminals do think about getting caught, these thoughts are easily dismissed in the offending moment. Similarly, Copes and Vieraitis' (2008) interviews with identity thieves found that most offenders simply did not think about the possibility of getting caught. For the most part, they were extremely confident in their ability to avoid detection and capture. Moreover, the few thieves who did consider the probability of arrest were able to put it out of their minds during the offending moment. In general, criminological research finds that most offenders give little thought to the potential legal consequences of their actions. That is not to say, however, that offenders never think of the possible legal sanctions, but that it is uncommon and more punitive sanctions are not likely to raise the "costs" appreciably.

One issue that originates from the aforementioned discussion of criminological research on offender decision making is whether the DP creates a general deterrent effect such that noncriminals avoid killing because of the potential consequences. Research on homicide offenders suggests that this scenario is highly unlikely. Most homicide offenders are not otherwise law-abiding citizens who killed during an act of passion or duress. According to Kleck and Kates (2001), most adult murderers have long histories of violence, felony records, and substance abuse. Wolfgang's classic 1958 study of homicide offenders found that 64.4% of his sample of homicide offenders had a previous arrest, 66% of whom had a record of crimes against the person. DeLisi's (2001) examination of career criminals found that murderers were significantly more likely to have arrests for violent index crimes, felony convictions, and prison sentences than other predatory offenders. Furthermore, DeLisi and Scherer (2006) found that the average murderer in their sample of 654 convicted and incarcerated homicide offenders from eight states had more than three prior felony convictions, more than one previous prison sentence, and approximately two probationary sentences. They also found that homicide offenders' arrest histories contained a mixture of violent and property index offenses as well as weapon and drug violations. Recent national statistics indicate that 65.5% of prisoners under a sentence of death had prior felony convictions, 8.4% of whom had prior homicide convictions (Bureau of Justice Statistics, 2007). As Elliott (1998: 1,085) notes, "life-threatening violence … is, in fact, largely restricted to a criminal class and embedded in a general pattern of criminal behavior." Thus, from this research, we know that most homicide offenders are not otherwise law-abiding citizens. Instead, they are repeat offenders who have been undeterred by prior arrests, convictions, and/or sentences of punishment.

The Nature of Homicide

Our knowledge of offenders' assessments of the potential consequences of their crimes is predominantly based on interviews with active or incarcerated offenders who have engaged in street property crimes (for exceptions, see Copes and Vieraitis, 2008; Shover, Coffey, and Sanders, 2004). Although this research suggests that property offenders do engage in at least some planning (e.g., target selection), it also suggests that these individuals typically do not evaluate the risks of legal sanctions. In light of these findings, it is probably even less likely that violent offenders do (Piquero and Rengert, 1999). Nonetheless, although we might conclude that property offenders are unlikely to be deterred by legal sanctions, the issue relevant to DP research is whether and to what extent these findings can be applied to homicide offenders.

Nagin and Pogarsky (2004) found that property and violent crime may be associated with different approaches to the consideration of future consequences. Property offenders tended to discount or deliberately devalue future consequences, whereas violent offenders tended not to consider future consequences at all. This finding is not surprising in light of the fact that violent offenders, including homicide offenders, are more likely in a state of emotional duress at the time of the crime and are unlikely to be thinking of the costs of their actions. Moreover,

many homicide offenses are committed during the commission of another felony (e.g., rob-bery). If committing a less serious felony is the goal of the offender, then it is unlikely they are simultaneously weighing the possibility of a death sentence for a murder they were not originally planning to commit.

In sum, for the DP to serve as a deterrent to homicide, potential offenders need at least to consider the possibility that they may be caught and that the probability of this occurring is greater than any benefit they may receive, whether monetary or psychologically, for killing someone. A substantial body of criminological research on offenders' decision-making processes and the dynamics of homicide events contradicts the supposition that criminals spend any notable amount of time considering the deleterious consequences of their actions. Instead, ethnographic research suggests these individuals are more likely to focus on the potential gains of their crimes rather than on the costs. Offenders who do consider the potential legal sanctions can easily minimize their fear of arrest and punishment and essentially nullify any deterrent effects of sanctions (Jacobs, 1999; Hochstetler and Copes, 2006; Tunnell, 1992). In addition, research by Kleck et al. (2005) demonstrated that criminals' and noncriminals' perceptions of the certainty, severity, and swiftness of punishment show little correspondence with actual punishment levels in their communities. Moreover, criminals were less accurate in their estima-tions than noncriminals. Thus, if neither criminals nor noncriminals (i.e., potential criminals) have little accurate information about legal sanctions, then the deterrent effect of the DP is unlikely to be substantial.

Although we acknowledge that economists' contributions toward criminological research has increased the methodological rigor of studies examining important crime policy issues such as the death penalty, the largely consumer-driven theoretical orientation of economics has promoted a simplistic description and has proposed a solution to the crime problem in the United States. By collectively ignoring the advances made in criminological theory on deterrence and offenders, econometric models lack critical empirical grounding and ignore key variables that may play into the death penalty-deterrence equation. Although claims of absolute and consistent deterrent effects might make for a great sound bite and fit nicely into the political agendas of lawmakers who endorse punitive crime-control policies and claim to be "tough on crime" (Blumstein, 1997; Currie, 2004; Lab, 2004), our research suggests that homicide rates are not influenced by any number of death penalty measures (i.e., the presence of a statute, the risk of execution, and the numbers of executions) or policy-related variables (i.e., 3X statutes, right-to-carry laws, crack index, or worsening prison conditions).

Most criminal behavior is not preceded by a simple calculus whereby the threat of pun-ishment or fear of an unlikely state-imposed death sentence will make offenders repent and reevaluate their antisocial lifestyles. Some scholars have argued persuasively that recent DP research by economists has "used econometric sophistication to silence debate rather than en-

lighten policymakers" (Donohue and Wolfers, 2005: 842). Thus, although the use of technically sophisticated methodology might be appealing, "intuitive plausibility should always be preferred in the realm of real-world policy. Unfortunately, the history of the DP debate is replete with examples of plausibility being sacrificed on the altar of sophistication" (Donohue and Wolfers, 2005: 842). Indeed, the assumption that offenders are average people who conduct a rational cost–benefit analysis prior to committing serious crimes, and who therefore can be deterred from committing capital-eligible offenses, is highly unlikely and inconsistent with much of the research on criminal decision making. Again, although some offenders may follow Becker's model of rational decision making and factor in the potential legal sanctions of their actions, it is likely a small portion of offenders.

In sum, our finding of no deterrent effect of the DP on homicide suggests the risk of execution does not enhance the level of deterrence. Therefore, we conclude that although policy makers and the public may continue to support the use of the death penalty based on retribution, religious grounds, or other justifications, defending its use based on deterrence is inconsistent with our findings. At a minimum, policy makers should refrain from justifying its use by claiming that it is a deterrent to homicide and explore less costly, more effective ways of addressing crime. In addition, research on the DP or any major policy issue that makes assumptions on how offenders consider the costs and benefits of their actions should be grounded theoretically and empirically. Toward this end, criminologists have an important role to play in the newly reignited debate over the deterrent impact of the death penalty.

References

Austin, James and John Irwin. 2001. *It's about time: America's imprisonment binge*. Belmont, CA: Wadsworth.

Ayres, Ian and John J. Donohue III. 2003. Shooting down the "more guns less crime" hypothesis. *Stanford Law Review*, 55: 1193–1312.

Ayres, Ian and John J. Donohue III. 2009a. More guns, less crime fails again: The latest evidence from 1977 to 2006. *Econ Journal Watch*, 6: 218–233.

Ayres, Ian and John J. Donohue III. 2009b. Yet another refutation of the more guns, less crime hypothesis—With some help from Moody and Marvell. *Econ Journal Watch*, 6: 35–59

Bailey, William. 1983. Disaggregation in deterrence and death penalty research: The case of murder in Chicago. *Journal of Criminal Law and Criminology*, 74: 827–859.

Baldus, David C. and James W.L. Cole. 1975. A comparison of the work of Thorsten Sellin and Isaac Ehrlich on the deterrent effect of capital punishment. *Yale Law Journal*, 85: 170–186.

Beck, Nathaniel and Jonathan N. Katz. 1995. What to do (and not to do) with time-series cross-section data. *American Political Science Review*, 89: 634–647.

Beck, Nathaniel and Jonathan N. Katz. 1996. Nuisance vs. substance: Specifying and estimating time-series–cross-section models. *Political Analysis,* 6: 1–36.

Beck, Nathaniel and Jonathan N. Katz. 2004. *Time-series-cross-section issues: Dynamics*. Unpublished manuscript.

Becker, Gary S. 1967. Crime and punishment: An economic approach. *Journal of Political Economy*, 76: 169–217.

Bedau, Hugo Adam. 1997. *The death penalty in America: Current controversies*. New York: Oxford University Press.

Bennett, Trevor and Richard Wright. 1984. *Burglars on burglary: Prevention and the offender*. Aldershot, UK: Gower.

Blumstein, Alfred. 1995. Youth violence, guns, and the illicit drug industry. *Journal of Criminal Law and Criminology*, 86: 10–36.

Blumstein, Alfred. 1997. Interaction of criminological research and public policy. *Journal of Quantitative Criminology*, 12: 349–361.

Blumstein, Alfred, Jacqueline Cohen, and Daniel S. Nagin. 1978. *Deterrence and incapacitation: Estimating the effects of criminal sanctions on crime rates*. Washington, DC: National Academy of Sciences.

Bourgois, Philippe. 1995. *In search of respect: Selling crack in el Barrio*. Cambridge, UK: Cambridge University Press.

Bowers, William J. and Glenn L. Pierce. 1975. The illusion of deterrence in Isaac Ehrlich's research on capital punishment. *Yale Law Journal*, 85: 187–208.

Brace, Paul and Brent D. Boyea. 2008. State public opinion, the death penalty, and the practice of electing judges. *American Journal of Political Science*, 52: 360–372.

Bureau of Justice Statistics. 2005. *Capital punishment in the United States, 1973–2005*. Ann Arbor, MI: Interuniversity Consortium for Political and Social Research.

Bureau of Justice Statistics. 2006. *Criminal victimization in the United States, 2005—Statistical tables*. Retrieved November 1, 2008 from ojp.usdoj.gov/bjs.

Bureau of Justice Statistics. 2007. *Capital punishment, 2006—Statistical tables*. Retrieved March 8, 2008 from ojp.usdoj.gov/bjs/pub/html/cp/2006/tables/cp06st08.htm

Bureau of Justice Statistics. 2008. *Deaths in custody reporting program*. Retrieved October 29, 2008 from ojp.usdoj.gov/bjs/dcrp/prisonindex.htm.

Campbell, Donald. T. and Julian Stanley. 1963. *Experimental and quasi-experimental designs for research*. Boston, MA: Houghton Mifflin Company.

Cherbonneau, Michael and Heith Copes. 2006. Drive it like you stole it: Auto thieves and the illusion of normalcy. *British Journal of Criminology*, 46: 193–211.

Cloninger, Dale O. and Roberto Marchesini, 2006. Execution moratoriums, commutations, and deterrence: The case of Illinois. *Applied Economics*, 38: 967–973.

Cochran, John K. and Mitchell B. Chamlin. 2000. Deterrence and brutalization: The dual effects of executions. *Justice Quarterly*, 17: 685–706.

Copes, Heith and Lynne M. Vieraitis. 2008. The risks, rewards, and strategies of stealing identities. In (Megan M. McNally and Graeme R. Newman, eds.), *Perspectives on identity theft. Crime prevention studies*, Vol. 23. Monsey, NY: Criminal Justice Press.

Cornish, Derek B. and Ronald V. Clarke. 1986. *The reasoning criminal: Rational choice perspectives on offending*. New York: Springer-Verlag.

Cromwell, Paul and James N. Olson. 2004. *Breaking and entering: Burglars on burglary*. Belmont, CA: Wadsworth.

Currie, Elliott. 2004. Against marginality: Arguments for a public criminology. *Theoretical Criminology*, 11: 175–190.

Dann, Robert. 1935. *The deterrent effect of capital punishment*. Philadelphia, PA: Friends Social Services Series.

DeLisi, Matt. 2001. Extreme career criminals. *American Journal of Criminal Justice*, 25: 239–252.

DeLisi, Matt and Aaron M. Scherer. 2006. Multiple homicide offenders: Offense characteristics, social correlates, and criminal careers. *Criminal Justice and Behavior*, 33: 367–391.

Dezhbakhsh, Hashem and Paul H. Rubin. 2007. *From the "econometrics of capital punishment" to the "capital punishment" of econometrics: On the use and abuse of sensitivity analysis*. Department of Economics Working Papers, Atlanta GA, Emory University.

Dezhbakhsh, Hashem, Paul H. Rubin, and Joanna M. Shepherd. 2003. Does capital punishment have a deterrent effect? New evidence from postmoratorium panel data. *American Law and Economics Review*, 5: 344–376.

Dezhbakhsh, Hashem and Joanna M. Shepherd. 2006. The deterrent effect of capital punishment: Evidence from a "judicial experiment." *Economic Inquiry*, 44: 512–535.

Donohue, John J. and Justin Wolfers. 2005. Uses and abuses of empirical evidence in the death penalty debate. *Stanford Law Review*, 58: 791–846.

Donohue, John J. and Justin Wolfers. 2006. The death penalty: No evidence for deterrence. *Economists' Voice*, April: 1–6.

Dunning, Thad. 2005. *Improving causal inference: Strengths and limitations of natural experiments*. Paper presented at the American Political Science Association's annual meeting. Washington, DC.

Ehrlich, Isaac. 1975. The deterrent effect of capital punishment: A question of life and death. *American Economic Review*, 65: 397–417.

Ekelund, Robert B., Jr., John D. Jackson, Rand W. Ressler, and Robert D. Tollison. 2006. Marginal deterrence and multiple murders. *Southern Economic Journal*, 72: 521–541.

Elliott, Delbert S. 1998. Life-threatening violence is primarily a crime problem: A focus on prevention. *University of Colorado Law Review*, 69: 1081–1098.

Fagan, Jeffrey. 2005a. *Deterrence and the death penalty: A critical review of new evidence*. Testimony to the New York State Assembly Standing Committee on Codes, Assembly Standing Committee on Judiciary and Assembly Standing Committee on Correction. Hearings on the Future of Capital Punishment in the State of New York. January 21, 2005.

Fagan, Jeffrey. 2005b. *Public policy choices on deterrence and the death penalty: A critical review of the new evidence*. Testimony before the Joint Committee on the Judiciary of the Massachusetts Legislature on House Bill 3834. An Act Reinstating Capital Punishment in the Commonwealth. July 14, 2005.

Fagan, Jeffrey. 2006. Death and deterrence redux: Science, law, and causal reasoning on capital punishment. *Ohio State Journal of Criminal Law*, 4: 255–320.

Fagan, Jeffrey, Franklin E. Zimring, and Amanda Geller. 2006. Capital punishment and capital murder: Market share and the deterrent effects of the death penalty. *Texas Law Review*, 84: 1803–1867.

Fryer, Roland G. Jr., Paul S. Heaton, Steven D. Levitt, and Kevin M. Murphy. 2005. *Measuring the impact of crack cocaine*. Working Paper 11318. National Bureau of Economic Research, Washington, DC.

Hochstetler, Andy and Heith Copes. 2006. Managing fear to commit felony theft. In (Paul Cromwell, ed.), *In their own words: Criminals on crime*. Los Angeles, CA: Roxbury.

Huber, Gregory A. and Sanford C. Gordon. 2004. Acccountability and coercion: Is justice blind when it runs for office? *American Journal of Political Science*, 48: 247–263.

Im, Kyung S., M. Hashem Pesaran, and Yongcheol Shin. 2003. Testing for unit roots in heterogeneous panels. *Journal of Econometrics*, 115: 53–74.

Jacobs, Bruce A. 1999. *Dealing crack: The social world of streetcorner selling*. Boston, MA: Northeastern University Press.

Katz, Lawrence, Steven D. Levitt, and Ellen Shustorovich. 2003. Prison conditions, capital punishment, and deterrence. *American Law and Economics Review*, 5: 318–343.

Kleck, Gary. 1979. Capital punishment, gun ownership, and homicide. *American Journal of Sociology*, 84: 882–910.

Kleck, Gary and Don B. Kates. 2001. *Armed: New perspectives on gun control*. Amherst, NY: Prometheus Books.

Kleck, Gary, Brion Sever, Spencer Li, and Marc Gertz. 2005. The missing link in general deterrence research. *Criminology*, 43: 623–659.

Kovandzic, Tomislav V., Thomas B. Marvell, and Lynne M. Vieraitis. 2005. Right-to-carry concealed handguns and homicide: Evidence from city panel data. *Homicide Studies*, 9: 292–323.

Kovandzic, Tomislav V., Mark E. Schaffer, and Gary Kleck. 2005. *Gun prevalence, homicide rates, and causality: A GMM approach to endogeneity bias*. CEPR Discussion Paper, 5357.

Kovandzic, Tomislav V., John J. Sloan III, and Lynne M. Vieraitis. 2002. The unintended consequences of politically popular sentencing policy: The homicide promoting effects of "three strikes laws" in U.S. cities (1980–1999). *Criminology & Public Policy*, 1: 399–424.

Lab, Steven. 2004. Crime prevention, politics, and the art of going nowhere fast. *Justice Quarterly*, 21: 681–692.

Lamperti, John. 2008. *Does capital punishment deter murder? A brief look at the evidence*. Retrieved October 8, 2009 from math.dartmouth.edu/~lamperti/capitalpunishment. pdf.

Levitt, Steven D. 1996. The effect of prison population size on crime rates: Evidence from prison overcrowding litigation. *Quarterly Journal of Economics*, 111: 319–351.

Levitt, Steven D. 1999. The limited role of changing age structure in explaining aggregate crime rates. *Criminology*, 37: 581–598.

Lott, John R., Jr. 2000. *More guns less crime*. Chicago, IL: University of Chicago Press.

Lott, John R., Jr. and David B. Mustard. 1997. Crime, deterrence, and right-to-carry concealed handguns. *Journal of Legal Studies*, 26: 1–68.

Marvell, Thomas B. 2001. The impact of banning juvenile gun possession. *Journal of Law and Economics*, 44: 691–714.

Marvell, Thomas B. and Carlisle E. Moody. 1994. Prison population growth and crime reduction. *Journal of Quantitative Criminology*, 7: 237–273.

Marvell, Thomas B. and Carlisle E. Moody. 1996. Specification problems, police levels, and crime rates. *Criminology*, 34: 609–646.

Marvell, Thomas B. and Carlisle E. Moody. 2001. The lethal effects of three-strikes laws. *Journal of Legal Studies*, 30: 89–106.

Mocan, H. Naci and R. Kaj Gittings. 2003. Getting off death row: Commuted sentences and the deterrent effect of capital punishment. *Journal of Law and Economics*, 46: 453–478.

Moody, Carlisle E. 2005. *Is crime a random walk?* Unpublished manuscript.

Nagin, Daniel S. and Greg Pogarsky. 2004. Time and punishment: Delayed consequences and criminal behavior. *Journal of Quantitative Criminology*, 20: 295–317.

National Institute of Justice. 2003. *2000 Arrestee drug abuse monitoring: Annual report*. Retrieved November 1, 2008 from ncjrs.gov/pdffiles1/nij/193013.pdf.

Peterson, Ruth D. and William Bailey. 1991. Felony murder and capital punishment: An examination of the deterrence question. *Criminology*, 29: 367–395.

Piquero, Alex R. and George F. Rengert. 1999. Studying deterrence with active residential burglars. *Justice Quarterly*, 16: 451–471.

Rao, Potluri. 1971. Some notes on misspecification in multiple regressions. *The American Statistician*, 25: 37–39.

Rengert, George F. and John Wasilchick. 1985. *Suburban burglary*. Springfield, IL: Charles C. Thomas.

Savitz, Leonard. 1958. A study in capital punishment. *Journal of Criminal Law, Criminology, and Police Science*, 49: 338–341.

Sellin, Thorsten. 1959. *The death penalty*. Philadelphia, PA: American Law Institute.

Shepherd, Joanna M. 2002. Fear of the first strike: The full deterrent effect of California's two- and three-strikes legislation. *Journal of Legal Studies*, 31: 159–201.

Shepherd, Joanna M. 2004. Murders of passion, execution delays, and the deterrence of capital punishment. *Journal of Legal Studies*, 33: 283–321.

Shepherd, Joanna M. 2005. Deterrence versus brutalization: Capital punishment's differing impacts among states. *Michigan Law Review*, 104: 203–256.

Shover, Neal. 1996. *Great pretenders: Pursuits and careers of persistent thieves*. Boulder, CO: Westview.

Shover, Neal, Glenn Coffey, and Clinton Sanders. 2004. Dialing for dollars: Opportunities, justifications, and telemarketing fraud. *Qualitative Sociology*, 27: 59–75.

Shover, Neal and David Honaker. 1992. The socially bounded decision making of persistent property offenders. *Howard Journal of Criminal Justice*, 31: 276–293.

Staiger, Douglas and James H. Stock. 1997. Instrumental variables regression with weak instruments. *Econometrica*, 65: 557–586.

Stolzenberg, Lisa and Stuart J. D'Alessio. 2004. Capital punishment, execution publicity, and murder in Houston, Texas. *Journal of Criminal Law and Criminology*, 94: 351–380.

Tunnell, Kenneth D. 1992. *Choosing crime: The criminal calculus of property offenders*. Chicago, IL: Nelson-Hall.

Van den Haag, Ernest. 1969. On deterrence and the death penalty. *Journal of Criminal Law, Criminology, and Police Science*, 60: 141–147.

Wolfgang, Marvin E. 1958. *Patterns in criminal homicide*. Philadelphia: University of Pennsylvania.

Wooldridge, Jeffrey M. 2001. *Econometric analysis of cross section and panel data*. Cambridge, MA: MIT Press.

Wooldridge, Jeffrey M. 2005. *Introductory econometrics: A modern approach*, 3rd Edition. Cincinnati, OH: South-Western College Publishing.

Wright, Richard and Scott H. Decker. 1994. *Burglars on the job: Streetlife and residential break-ins*. Boston, MA: Northeastern University Press.

Wright, Richard and Scott H. Decker. 1997. *Armed robbers in action: Stickups and street culture*. Boston, MA: Northeastern University Press.

Yang, Bijou and David Lester. 2008. The deterrent effect of executions: A meta-analysis thirty years after Ehrlich. *Journal of Criminal Justice*, 36: 453–460.

Zimmerman, Paul R. 2004. State executions, deterrence, and the incidence of murder. *Journal of Applied Economics*, 7: 163–193.

Zimmerman, Paul R. 2006. Estimates of the deterrent effect of alternative execution methods in the United States: 1978–2000. *American Journal of Economics and Sociology*, 65: 909–941.

Case Cited

Furman v. Georgia, 408 U.S. 238 (1972).

Gregg v. Georgia, 428 U.S. 153 (1976).

Tomislav V. Kovandzic, Ph.D., is an associate professor of criminology at the University of Texas at Dallas. His research interests include gun-related violence, criminal justice policy, and inequality and violent crime. He received his Ph.D. in criminology from Florida State University in 1999.

Lynne M. Vieraitis, Ph.D., is an associate professor of criminology at the University of Texas at Dallas. Her research focuses on studying the impact of criminal justice policy on crime rates, inequality and violence, and identity theft.

Denise Paquette Boots, Ph.D., is an assistant professor of criminology at the University of Texas at Dallas. Her most recent research projects focus on mental health problems as a predictor of violence over the life course, the gender gap and capital punishment, and the social learning and intergenerational transmission of violence mechanisms within domestic batterers.

POLICY ESSAY

DETERRENCE AND EXECUTIONS

Can't tell

Comments on "Does the death penalty save lives?"

Richard Berk
University of Pennsylvania

Tomislav V. Kovandzic, Lynne M. Vieraitis, and Denise Paquette Boots (2009) offer us another round in the dueling econometrics of the death penalty. They bring to the dispute the most recent data readily available, substantial theoretical sophistication, and an encyclopedic application of standard econometrics for panel data. They find no evidence for deterrent effects. We have probably not, however, heard the end of this issue. With each new year, another year's worth of data become available. There will be variation over time in death penalty practices and in homicide patterns. The political process will continue to seek findings to its liking. And researchers with sunk costs in the death penalty debates will have strong incentives to stay the course.

Yet, we are unlikely ever to know what impact on crime the death penalty may have had. Kovandzic et al.'s (2009) study is my sense of where we were 20 years ago, where we are now, and where we will be 20 years in the future. Therefore, as long as we have some researchers who claim to have found deterrent effects for the death penalty, it is good that we have other researchers who claim not to have found deterrent effects. We always will have advocates who are looking for simple and congenial answers to the most difficult empirical questions. Therefore, a message that simple answers do not exist, when in fact none do exist, is healthy.

That said, I am fundamentally skeptical about the research tradition this work represents. My comments, therefore, will focus on why "can't tell" is where we now stand and where we will be standing in the future. I also will suggest more promising alternatives and briefly conclude with lessons for policy. The discussion is necessarily brief, but extensive references are provided for those readers who want to consider the issues in more depth.

Direct correspondence to Richard Berk, Department of Criminology, University of Pennsylvania, 483 McNeil Building, 3718 Locust Walk, Philadelphia, PA 19104 (e-mail: berkr@sas.upenn.edu).

Well-Known Problems with the Conventional Econometric Paradigm

Growing out of economic research undertaken 60 years ago and becoming the dominant sta-tistical formulation within economics, political science, sociology, and criminology during the 1970s, the econometric paradigm—which can found in any elementary econometrics textbook (Cameron and Trivedi, 2005: 3–10; Pindyck and Rubinfeld, 1981: xiiixvii)—has the following features. It depends fundamentally on a model, which is a simplified theory of how the data were generated. The canonical model can be represented as

$$y_i = f(X_i) + \varepsilon_i \qquad\qquad\qquad (1)$$

where i = 1, 2, … *N* observations, y_i is the quantitative response variable, X_i is a matrix that includes all explanatory variables (and usually a column of 1s for the constant), and ε_i is an unobservable disturbance term that behaves as if each disturbance is drawn independently of one another and X_i from a single distribution with a mean of 0.0 and a common variance. The $f(X_i)$ needs to be specified before looking at the data. The canonical model can be extended to encompass (1) more than one equation, (2) response variable(s) that are categorical or ordinal, (3) data that can be arrayed in time and space, and (4) less-restrictive disturbance terms proper-ties. However, the basic features remain.

In the absence of a randomized experiment, one bedrock assumption exists that researchers cannot avoid no matter how elaborate the econometrics employed. To obtain proper estimates of the key regression parameters, the disturbances must be uncorrelated with the regressors—at least in large samples. At a minimum, this process requires that (1) one has in X_i all regressors required by the "correct" theory (or that one can persuasively argue that the omitted regressors captured by ε_i are uncorrelated with regressors in the model); (2) that all regressors and the response variable(s) are measured well;[1] and (3) that one knows and can specify the correct $f(X_i)$. In practice, it is impossible to know whether these requirements are met (Berk, 2003: Section 9.6; Freedman, 2005: Section 8.6), especially with observational data of the kind used in death-penalty deterrence research (Berk, 2005), and ample reasons abound for enormous skepticism. For example, if you read the fine print in the primary sources, then you will learn that you cannot test your way to the correct model unless you are prepared to make several heroic assumptions about what is not being tested. As the literature on deterrence and the death penalty illustrates well, endless debate of the "my-model-is-better-than-your-model" variety can follow from these problems. Now, we have at least 40 years of sophisticated criticism of the econometric paradigm (Berk, 2003; Box, 1976; Breiman, 2001; de Leeuw, 1994; Holland, 1986; Leamer, 1978; Lieberson, 1985; Rubin, 1986; Freedman, 1987, 2005; Kish, 1987; Pearl, 2000), and some of the most interesting recent criticism comes for the leading econometricians (Heckman, 2000; Imbens, 2009; Manski, 1995). I think a fair response to econometrician Ed

---

1. The response variable should be measured with no bias and, ideally, with little random measurement error. The regressors must be measured with no bias and with no random measurement whatsoever—at least in theory. It is important to appreciate that even random measurement error (i.e., independent errors with a mean 0.0 uncorrelated with the true score) is not permitted formally for the regressors.

Leamer's early plea (1983) to "take the con out of econometrics" is that the con is built in. We need other approaches.

Alternatives for Observational Data

If traditional econometric methods can be a dead end for most observational data sets, then what are the alternatives? Elsewhere, this matter is addressed in depth (Berk, 2003: Ch. 11; Berk, Sherman, Barnes, Kurtz, and Ahlman, 2009; Breiman, 2001; Gifi, 1990; Imbens, 2004). A few highlights will have to suffice here.

True Experiments and Quasi-Experiments

When traditional econometric procedures are applied to observational data, they risk being compromised by weak research designs. Since the mid-1960s, it has been apparent that a wide range of designs can improve on conventional observational studies (Campbell and Stanley, 1963). If causal inference is the top priority, then randomized experiments are the ideal. When they are not practical, several strong quasi-experimental designs often will perform nearly as well (Berk, Barnes, Ahlman, and Kurtz, 2010; Cook, Shadish, and Wong, 2008). Indeed, regression continuity designs are one topic on which econometricians have recently made important contributions (Imbens and Lemieux, 2008). In the absence of either a randomized experiment or a strong quasi-experiment, arriving at credible estimates of causal effects through econometric modeling can be difficult.

Model-Lite Causal Analysis

When one has no choice but to work with observational data, as is the case here, approaches to causal inference are available that are much less model dependent, more transparent, and more credible. Morgan and Winship (2007) provide an excellent overview that builds on the work of William Coschran, Donald Rubin, Paul Rosenbaum, Judea Pearl, and many others. The basic strategy is to conceptualize observational data within a randomized experiment framework that is conditional on a set of covariates. This strategy has several benefits (Rubin, 2008). A reliance on the Neyman–Rubin casual model makes a critical distinction between one or more causal variables and one or more covariates, or nuisance variables. Casual inference is of no concern for the covariates. This point permits a much clearer rendering of appropriate counterfactuals than death-penalty deterrence research has achieved to date. It also fosters analysis strategies that do not have to depend as heavily on untestable assumptions. For example, because one does not care about the causal impact of the covariates, matching or stratifying on those covariates can be a useful approach (Rosenbaum, 2002). Finally, this perspective helps prevent data snooping (Freedman, 2005: Section 4.9). The response variable is ignored while efforts to achieve balance in the data are undertaken. One cannot capitalize on chance that leads to overfitting. It also precludes model selection that is recognized increasingly as a danger to causal inference and estimation far more fundamentally (Leeb and Pötscher, 2005).

"Mere" Description

A long tradition in statistics and some social sciences favors quantitative description. Among the preferred tools are multivariate statistics (Anderson, 1958; Gifi, 1990), exploratory data analysis (Diaconis, 1985; Tukey, 1977), dynamic graphics (Cook and Swayne, 2007), and machine/statistical learning (Hastie, Tibshirani, and Friedman, 2009; Lohr, 2009). The intent is to find systematic patterns in data without worrying about cause and effect. With modern algorithms and powerful computers, the search possibilities virtually are limitless. Some researchers see this technology as a precursor to explicit model building. Others see it as a substitute for model building, especially when credible model building is not practical. These approaches are finding their way into criminology in several interesting applications. GIS studies of crime are one common illustration (Chainey and Ratcliffe, 2005; Groff and La Vigne, 2001). Others include inductively developed crime forecasting models (Berk, Barnes, Ahlman, and Kurtz, 2010) and capturing crime dynamics in time and space (Berk and McDonald, 2009).

Conclusions

Responding several years ago to possibly the same set of studies that motivate the Kovandzic et al.'s (2009) work, I wrote (Berk, 2005: 328): "[My] analyses reported here are hardly exhaustive and are perhaps affected by [my] misunderstandings about the data provided, or by errors in the data themselves. Nevertheless, the results raise serious questions about whether anything useful about the deterrent value of the death penalty can ever be learned from an observational study with the data that are likely to be available." The panel data have been made more current, and a wider rage of econometric procedures have been brought to bear. But the observational design still is weak and the measurement problems in the data essentially remain, as do all the questions that naturally follow from trying to build a credible casual model.

I stand by my earlier conclusions. Nevertheless, Kovandzic et al. (2009) have done a great service by showing once again that conventional econometric modeling can be a fragile enterprise. They also have helped to debunk claims that the death penalty has demonstrable deterrent value and, possibly more important, they perhaps have motivated some researchers to consider alternative empirical approaches. Three lessons for evidence-based policy follow. First, no credible evidence exists that the death penalty, as implemented in the United States since 1979, has any deterrent value. Second, no credible evidence exists to rule out any deterrent effects. Third, if evidence-based policy is to be more than another catch-phase, then one must be careful about what qualifies as evidence.

When causal effects are primary, evidence-based policy means implementing true experiments or strong quasi-experiments. When causal effects are important, but significant tradeoffs exist with other priorities, observational studies analyzed through well-executed matching methods can provide useful evidence. When estimates of casual effects are secondary, premature, or irrelevant, description alone can be instructive. Finally, judgment must be used to identify those empirical questions whose answers are beyond the data and the methods by which those data

can be analyzed. Trying to estimate the causal impact of the death penalty on crime is one good illustration. There are others.

References

Anderson, Theodore. 1958. *An introduction to multivariate statistical analysis*, 1st Edition. New York: Wiley.

Berk, Richard A. 2003. *Regression analysis: A constructive critique*. Newbury Park, CA: Sage.

Berk, Richard A. 2005. New claims about executions and general deterrence: De ja vu all over again. *Journal of Empirical Legal Studies*, 2: 303–330.

Berk, Richard A. 2009. What now? Some brief reflections on model-free data analysis. *International Econometric Review,* 1: 14–23.

Berk, Richard A., Geoffrey Barnes, Lindsay Ahlman, and Ellen Kurtz. 2010. When second best is good enough: A comparison between a true experiment and a regression discontinuity quasi-experiment. *Journal of Experimental Criminology*. In press.

Berk, Richard A. and John McDonald. 2009. The dynamics of crime regimes. *Criminology*. In press.

Berk, Richard A., Lawrence Sherman, Geoffrey Barnes, Ellen Kurtz, and Lindsay Ahlman. 2009. Forecasting murder within a population of probationers and parolees: A high stakes application of statistical forecasting. *Journal of the Royal Statistical Society*, 172: 191–211.

Box, George E. P. 1976. Science and statistics. *Journal of the American Statistical Association*, 71: 791–799.

Breiman, Leo. 2001. Statistical modeling: Two cultures (with discussion). *Statistical Science*, 16 :199–231.

Cameron, A. Colin and Pravin K. Trivedi. 2005. *Microeconomics: Methods and applications*. Cambridge, UK: Cambridge University Press.

Campbell, Donald T. and Julian Stanley. 1963. *Experimental and quasi-experimental designs for research*. New York: Wadsworth.

Chainey, Spencer and Jerry Ratcliffe. 2005. *GIS and crime mapping*. New York: Wiley.

Cook, Dianne and Deborah F. Swayne. 2007. *Interactive dynamic graphics and data analysis*. New York: Springer.

Cook, Thomas D., William R. Shadish, and Vivian C. Wong 2008. Three conditions under which experiments and observational studies produce comparable causal estimates: new findings from within-study comparisons. *Journal of Policy Analysis and Management*, 727: 724–750.

de Leeuw, Jan. 1994. Statistics and the sciences, in (I. Borg and P.P. Mohler, eds.), *Trends and perspectives in empirical social science*. New York: Walter de Gruyter.

Diaconis, Persi. 1985. Theories of data analysis: From magical thinking through classical statistics. In (D.C. Hoaglin, F. Mosteller, and J. Tukey, eds.), *Exploring data tables, trends, and shapes*. New York: Wiley.

Freedman, David A. 1987. As others see us: A case study in path analysis (with discussion). *Journal of Educational Statistics*, 12: 101–223.

Freedman, David A. 2005. *Statistical models: Theory and practice*. Cambridge, UK: Cambridge University Press.

Gifi, Albert. 1990. *Nonlinear multivariate analysis*. New York: Wiley.

Groff, Elizabeth and Nancy G. La Vigne. 2002. Forecasting the future of predictive crime mapping. In (N. Tilley, ed.), *Analysis for crime prevention*. Monsey, NY: Criminal Justice Press.

Hastie, Trevor, Robert Tibshirani, and Jerome Friedman. 2009. *The elements of statistical learning*, 2nd Edition. New York: Springer.

Heckman, James J. 2000. Causal parameters and policy analysis in economics: A twentieth century retrospective. *Quarterly Journal of Economics*, 115: 45–97.

Holland, Pail W. 1986. Statistics and causal inference. *Journal of the American Statistical Association*, 81: 945–960.

Imbens, Guido W. 2004. Nonparametric estimation of average treatment effects under exogeneity: A review. *Review of Economics and Statistics*, 86: 4–29.

Imbens, Guido W. 2009. *Better late than nothing: Some comments on Deaton (2009) and Heckman and Urzua (2009)*. Working paper, Department of Economics, Harvard University, Cambridge, MA.

Imbens, Guido W. and Thomas Lemieux. 2008. Regression continuity designs: A guide to practice. *Journal of Econometrics*, 142: 615–635.

Kish, Leslie. 1987. *Statistical design for research*. New York: Wiley.

Kovandzic, Tomislav V., Lynne M. Vieraitis, and Denise Paquette Boots. 2009. Does the death penalty save lives? New evidence from state panel data, 1977 to 2006. *Criminology & Public Policy*. This issue.

Leamer, Edward E. 1978. *Specification searches: Ad hoc inference with non- experimental data*. New York: Wiley.

Leamer, Edward E. 1983. Let's take the con of econometrics. *American Economics Review*, 73: 31–43.

Leeb, Hannes and Benedikt M. Pötscher. 2005. Model selection and inference: Facts and fiction. *Econometric Theory*, 21: 21–59.

Lieberson, Stanley. 1985. *Making it count: The improvement of social research and theory*. Berkeley: University of California Press.

Lohr, Steven. 2009, August 5. For today's graduate, just one word: Statistics. *New York Times*.

Manski, Charles. 1995. *Identification problems in the social sciences*. Cambridge, UK: Cambridge University Press.

Morgan, Stephen L. and Christopher Winship. 2007. *Counterfactuals and causal inference: Methods and principles for social research*. Cambridge, UK: Cambridge University Press.

Pearl, Judea. 2000. *Causality: Models, reasoning and inference*. Cambridge, UK: Cambridge University Press.

Pindyck, Robert S. and Daniel L. Rubinfeld. 1981. *Econometric models and economic forecasts*. New York: McGraw Hill.

Rosenbaum, Paul 2002. Observational studies, 2nd Edition. New York: Springer-Verlag.

Rubin, Donald B. 1986. Which ifs have causal answers? *Journal of the American Statistical Association*, 81: 961–962.

Rubin, Donald B. 2008. For objective causal inference, design trumps analysis. *Annals of Applied Statistics*, 2: 808–840.

Tukey, John. 1977. *Exploratory data analysis*. Reading, MA: Addison-Wesley.

Richard Berk is a professor in the departments of statistics and criminology at the University of Pennsylvania. He is an elected fellow of the American Association for the Advancement of Science, the American Statistical Association, as well as the Academy of Experimental Criminology and has been a member of the Committee on Applied and Theoretical Statistics of the National Research Council and the Social Science Research Council's Board of Directors. He has been awarded the Paul F. Lazarsfeld Award by the Methodology Section of the American Sociological Association. Berk works on various topics in applied statistics that include causal inference, statistical learning, and methods for evaluating social programs. Among his criminology applications are inmate classification and placement systems, forecasts of behavior while on parole or probation, the role of race in capital punishment, detecting violations of environmental regulations, claims that the death penalty serves as a general deterrent, and describing and forecasting short-term changes in urban crime patterns. Berk's latest book is *Statistical learning from a regression perspective* for the Springer Series in Statistics (2008).

Don't scrap the death penalty

Paul H. Rubin

Emory University

I have been asked to comment on the policy implications of "Does the death penalty save lives? New evidence from state panel data, 1977 To 2006," (Kovandzic, Vieraitis, and Boots, 2009, this issue). The article does not hesitate to draw its own policy conclusions. Kovandzic et al. conclude that because no evidence is provided for a deterrent effect, "policy makers should refrain from justifying its use by claiming that it is a deterrent to homicide and consider less costly, more effective ways of addressing crime." I will be equally blunt; I do not think that their evidence is sufficiently persuasive to draw the conclusions they do, and I am not in favor of abolishing the death penalty.

I will consider two questions about the death penalty. First, as execution is currently applied in the United States, does it actually serve as a deterrent? This question will require an examination of some assertions in the Kovandzic et al. (2009) article about the nature of criminals as well as about the predictions of economic theory. Second, if the death penalty is currently a deterrent, is it powerful enough so that existing statistical techniques can measure the deterrent effect? This question will entail an examination of the statistical analysis of the article.

Is Capital Punishment a Deterrent Theory?

In regard to the first question, at some level of certainty and speed, capital punishment could serve as a deterrent. Gordon Tullock has proposed the following thought experiment: Think of a criminal pointing a gun at a potential victim with a policeman standing behind the criminal. The policeman says, "You can shoot him, but if you do then I will shoot you." In this circumstance, the threat of death certainly would deter most or all potential murderers. We then can think of actual systems as reducing the probability and speed of execution. With each reduction, the level of deterrence also is reduced because fewer potential murderers will be deterred as the probability and speed of execution is reduced.

Moreover, any probability of execution will deter some potential murderers. In some circumstances, the decision of whether to commit murder is marginal. The criminal sometimes will be on the cusp of murder ("Do I shoot the person I have just robbed or let him go?" "Do I kill the rape victim?" "Will a severe beating punish my enemy enough, or should I kill him?" "Do I kill my spouse or seek a divorce?"), and for at least some of these potential murderers, the extra deterrent will be enough to lead to a decision not to commit murder.

Kovandzic et al. (2009) deny that deterrence is present. But their language allows for some deterrence. Kovandzic et al. downplay the possibility of deterrence but cannot eliminate it. For example:

> A substantial body of criminological research on offenders' decision-making processes and the dynamics of homicide events contradicts the supposition that criminals *spend any notable amount of time* considering the deleterious consequences of their actions. Instead, ethnographic research suggests these individuals *are more likely* to focus on the potential gains of their crimes rather than on the costs (2009, emphasis added).

Both of these statements and others throughout the discussion section (e.g., "essentially nullify any deterrent effects" or "the deterrent effect of the DP is unlikely to be substantial") show that Kovandzic et al. believe that at least some deterrence is associated with more severe punishments.

Moreover, Kovandzic et al.'s (2009) argument proves too much. Even if the threat of execution does not deter criminals, it still is the case that a large body of evidence shows that increasing the probability and the severity of punishment in general deters crime. (For a summary, see Eide, Rubin, and Shepherd, 2006.) If it is true as Kovandzic et al. state that "Evidence from interviews with active and captured offenders suggests they do not dwell on the potential consequences of their actions and thus rarely consider the possibility of arrest and imprisonment," then no punishment would deter any crime, and that clearly is not the case.

Indeed, deterrence historically has been an important issue in the capital punishment debate. Capital punishment is imposed infrequently in the United States. Evidence of the extent to which it is a deterrent is difficult to obtain, and some studies (which includes the study at issue here) find that it is not a deterrent. But even if capital punishment was not or could not be shown to be a deterrent, that still does not apply to other punishments (which are much more certain and less erratic than capital punishment). It is possible, and generally uncontested, that increased probability and severity of jail sentences for crime reduces the amount of crime, and even those individuals who do not believe the evidence for capital punishment is persuasive generally accept that punishment does deter crime. This fact is an important point to keep in mind.

Another issue also plagues criminological research. Criminologists study criminals. For many purposes, this study is useful. But when the issue being studied is deterrence, then clearly criminals are a highly biased sample, because they have been caught; if criminals avoid detection, then they escape the criminologists' lens. But more importantly, criminologists miss all those

individuals who actually are deterred. That is, when deterrence does succeed, some potential criminals do not become criminals and so are not examined by criminologists. Criminologists' sample, therefore, is unsuited for any inference about deterrence. With respect to the issue of deterrence, this point can be a crucial issue.

Kovandzic et al. (2009) realize this issue and try to avoid it. They argue that "most homicide offenders are not otherwise law-abiding citizens who killed during an act of passion or duress." However, their data belie this argument. Kovandzic et al. cite a study (Wolfgang, 1958), which showed that 64% of homicide offenders had a previous arrest and 42% (66% of the 64% with a previous arrest) had a record of crimes against the person. In other words, 36% of murders had no previous arrest and 58% had no history of crimes against the person. A more recent study, which they cite (Bureau of Justice Statistics, 2007), found that 65.5% of criminals with a death sentence had prior felony convictions, so 34.5% did not. Given these numbers, clearly all murderers are not hardened criminals, and it is plausible that some potential murderers (obviously not in their sample) were deterred by the threat of execution. These potential murderers would not show up in any study of convicted murders.

Finally, Kovandzic et al. (2009) misunderstand the economic arguments in regard to crime. It is true that Becker's (1967) model is a formal cost–benefit analysis of committing a crime. This model is what Kovandzic et al. (2009) criticize: "Again, although some offenders might follow Becker's model of rational decision making and factor in the potential legal sanctions of their actions, it is likely a small portion of offenders." But for deterrence not to be present, it is not necessary that offenders undertake accurate cost-benefit analysis any more than it is necessary for consumers to undertake complex calculations of marginal use for the law of demand to hold. Rather, behavior must respond to changes in prices or costs. If one jurisdiction executes more criminals than another or if a jurisdiction becomes more likely to execute a criminal, and criminals perceive the direction of change in probability, then deterrence can be increased. Caricatures of economics that require criminals to be calculators of pleasure and pain are no more accurate than similar caricatures in other areas of behavior.

They also overstate the arguments of economists. Economists generally believe that punishment deters crime; we do not believe that feasible levels of punishment can lead to a "solution to America's crime problem." Rather, economists who study the issue believe that deterrence is one tool that can be used to reduce crime. We have studied this tool more than others, because deterrence is where we have a comparative advantage, not because we believe that it is the only solution. Economists believe in the division of labor, and clearly, criminologists have a role in studying methods of reducing crime. But, as I show in the next section, criminologists might do better to use the tools in which they have a comparative advantage rather than attempting to use econometric tools in their research.

Econometric Issues

First, I must state that Kovandzic et al.'s (2009) conclusions go beyond their evidence. They implicitly conclude that they have written the definitive study on capital punishment and deterrence and have shown deterrence is not present so that, with respect to capital punishment, "policy makers should refrain from justifying its use by claiming that it is a deterrent to homicide." Yet, they themselves cite a meta-analysis of empirical studies of capital punishment (Yang and Lester, 2008). This meta-analysis found 104 refereed studies, of which 95 had data adequate for analysis. Of these 95, 66 found a deterrent effect. Thus, Kovandzic et al's study is only one more data point. As such, it is somewhat hubristic to claim that it is the definitive study and should itself be the basis for policy. Indeed, to make such a claim would require examining all 66 studies that found a deterrent (or at least the 15 panel data studies that found a deterrent) and showing why this particular study is preferable. In fact, only 10 studies are cited (Table 1), and Kovandzic et al. actually do not compare many of these findings with their results.

Kovandzic et al. (2009) assert that they have accounted for differences:

> The most likely explanation for the divergence between our largely null findings and studies reporting robust deterrent effects that result from increases in execution risk is the failure of the latter to (1) address adequately omitted variable bias by failing to include year dummies and or state-specific trends in the regression model, (2) adjust standard errors to correct for serial correlation, and (3) rely on IV estimation using invalid instruments to address phantom simultaneity effects of homicide on execution risk.

I examine each point. I mainly will rely on the article I know best, Dezhbakhsh, Rubin, and Shepherd (2003). This article is cited in Kovandzic et al. (2009) and so is presumably one they believe to be surpassed by their research.

First, consider omitted variables. Dezhbakhsh et al. (2003) use panel data and includes year and county (or state) dummies; Dezhbakhsh, et al. does not include state-specific time trends. However, Dezhbakhsh et al. also include measures of probability of arrest as well as the conditional probability of a death sentence; both of these variables are omitted in the Kovandzic et al. (2009) specification. Dezhbakhsh et al. include aggravated assault and robbery rates as well as NRA membership and racial composition; Kovandzic et al. does not include any of these variables, which generally are significant in our specification (In our robustness checks, we dropped the measures of assault and robbery, and some of our deterrence measures became insignificant.) Thus, their specification includes some variables we had omitted but omits some significant variables we had included. Moreover, no strong reason or evidence is given for including state-specific time trends. As shown in Table 4 of Kovandzic et al., when this trend is dropped, the results become negative and significant for executions and homicide. More generally, the specification generally suffers from overparametrization. This specificity reduces the estimation efficiency, which leads to insignificant coefficient estimates. Examples include using state fixed effects and year fixed effects along with state-specific time trends. The

overparametrization is reflected in Table 3 in which few variables are significant in regression equations with R^2 of 94%.

Kovandzic et al. (2009) claim that the second improvement is an adjustment for serial correlation. They assert that the deterrent finding in some other studies has been the product of serial correlation. However, they have neither tested nor established the presence of serial correlation in their data set, let alone in the other data sets critiqued. Thus, this assertion is without foundation.

Third, Kovandzic et al. (2009) do not use simultaneous equations or instrumental variables; they estimate only ordinary least-squares specifications. But their discussion implies that simultaneity likely is to be an issue. At the end, however, they neither apply the methods used by other authors to address this simultaneity issue, nor do they use any valid alternative approach to treat this problem. Instead, they use the inadequate and simplistic method of looking at the bivariate correlation between homicide rate and the number of death-row inmates. This method is only a (weak) test for one possible source of simultaneity, and Kovandzic et al. do not test for any others. In our debate with Donohue and Wolfers (2005) and Dezhbakhsh and Rubin (2009), a discussion takes place of the proper instruments for controlling simultaneity, but Kovandzic et al. completely ignore this discussion and simply disregard the possibility of simultaneity.

Finally, it is improper to claim that the Kovandzic et al. (2009) article is stronger than the Dezhbakhsh et al. (2003) article because they ultimately are noncomparable. In Dezhbakhsh et al., deterrence is measured six times. However, all measurements are conditional on arrest and on sentencing to death. That is not the case in Kovandzic et al. Moreover, only one measure used in Kovandzic et al. is the same as any of the measures in Dezhbakhsh et al. (2003). Dezhbakhsh et al. uses executions lagged by 6 years, executions leading by 6 years, and a 6-year moving average. (The 6-year period is based on average lag between sentencing and execution.) In three specifications, jurisdictions with no murders or death sentences were excluded. In the other three specifications, the most recent available values were used to fill in cases in which the value is zero. Although Kovandzic et al. use eight measures of execution, only one (the 6-year lagged execution measure) is the same as Dezhbakhsh et al., and this specification had the smallest measured effect. Indeed, in the model in which state (rather than county) data were used, this measure is insignificant, although the other five measures were significant. Because Kovandzic et al. uses state data, the fact that it uses our weakest measure (which we also found to be insignificant in state data) indicates that the results are noncomparable. Kovandzic et al. (2009) do not consider or test the specifications in which significant deterrence was found.

Other econometric issues are present in the Kovandzic et al. (2009) article as well. Many selections and remedies are ad hoc and at odds with sound econometric practice. For example, introduction of lag of the dependent variable is used as a method to correct for serial correlation. Despite using heteroskedastic robust standard errors, Kovandzic et al. argue that weighted least square needs to be used to correct for heteroskedasticity. This assertion is not evidence-based, and the solution also is adopted without testing to see whether this remedy deals with the issue

of assumed heteroskedasticity. The specification does not address the choice of the functional form adequately, which ignores double log, Poisson, and negative binomial regressions that some other death penalty studies have used to show robustness. Kovandzic et al.'s statement that "Any deterrent impact due to the presence of these laws is captured by the state fixed-effects variables" is incorrect and misleading. Such effects only can be captured by extending the sampling period to those years.

In sum, Kovandzic et al (2009) change the model specification, estimation method, as well as both the dependent and independent variables used by earlier death penalty studies that report deterrence, and they find no deterrence. This conclusion is not surprising to those scholars well versed in econometrics. Kovandzic et al., however, interpret such a divergence of results as validation of their findings. Their logic reads as follows: If our results are different from others, then ours must be the correct one. To prove their assertions, Kovandzic et al. instead should have established, with rigor, that their results are derived from more appropriate statistical models and must, therefore, be the correct one. Moreover, their statistical methods are unjustified and, at times, inappropriate. Their assertion about the lack of a deterrent effect is, therefore, unwarranted given their evidence.

Finally, Kovandzic et al. (2009) overinterpret their results. They do not find evidence of deterrence, but this does not mean no deterrence is present. As a recent op-ed by two participants (on opposing sides) in the debate says: "But the absence of evidence of deterrence should not be confused with evidence of absence." (Sunstein and Wolfers, 2008:A11). Statistical tests fail unless criminologists can show some effect with greater (90% or 95%) levels of confidence. But the failure to show effects does not indicate that they do not exist.

Conclusion

The Kovandzic et al. (2009) article is one of more than 100 refereed articles that examine the deterrent effect of capital punishment. Most of these articles find a statistically significant deterrent effect (Yang and Lester, 2008). Moreover, it would be incredible and a violation of the law of demand if the chance of execution did not deter at least some murders. The data are murky, and it is difficult to measure empirically the effects under consideration, but the weight of the evidence as well as the theoretical predictions both argue for deterrence, and econometrically flawed studies such as this article are insufficient to overthrow this presumption.

One final note: The overwhelming weight of academic opinion is opposed to capital punishment. Even articles that find a deterrent are hesitant about drawing policy conclusions. For example, Yang and Lester (2008: 459) explained: "Even if executions are shown to deter potential murderers, alternative strategies to reduce the murder rate may be more effective and more ethically acceptable." But professors live in safe neighborhoods, have relatively safe jobs and are not likely to be drug dealers. Most murders occur in poor neighborhoods and among relatively uneducated persons, often with risky lifestyles. An element of elitism may be present in academic recommendations for abolishing the death penalty, because others will bear the costs.

References

Becker, Gary S. 1967. Crime and punishment: An economic approach. *Journal of Political Economy*, 76: 169–217.

Bureau of Justice Statistics. 2007. *Capital punishment, 2006—Statistical tables*. Retrieved September 22, 2009 from ojp.usdoj.gov/bjs/pub/html/cp/2006/tables/cp06st08.htm.

Dezhbakhsh, Hashem and Paul H. Rubin. 2009. *From the "econometrics of capital punishment" to the "capital punishment" of econometrics: On the use and abuse of sensitivity analysis*. Emory University working paper.

Dezhbakhsh, Hashem, Paul H. Rubin, and Joanna Shepherd. 2003. Does capital punishment have a deterrent effect? New evidence from postmoratorium panel data. *American Law and Economics Review*, 5: 344–376.

Donohue, John J. and Justin Wolfers. 2005. Uses and abuses of empirical evidence in the death penalty debate. *Stanford Law Review*, 58: 791–846.

Eide, Erling, Paul H. Rubin, and Joanna Shepherd. 2006. The economics of crime. *Foundations and Trends in Microeconomics*, 2: 291–363.

Kovandzic, Tomislav V., Lynne M. Vieraitis, and Denise Paquette Boots. 2009. Does the death penalty save lives? New evidence from state panel data, 1977 to 2006. *Criminology & Public Policy*. This issue.

Sunstein, Cass R. and Justin Wolfers. 2008. A death penalty puzzle. *Washington Post*, June 30: A11.

Wolfgang, Marvin E. 1958. *Patterns in criminal homicide*. Philadelphia: University of Pennsylvania.

Yang, Bijou and David Lester. 2008. The deterrent effect of executions: A meta-analysis thirty years after Ehrlich. *Journal of Criminal Justice*, 36: 453–460.

Paul H. Rubin is Samuel Candler Dobbs Professor of Economics at Emory University in Atlanta and editor-in-chief of *Managerial and Decision Economics*. He is a Fellow of the Public Choice Society and former Vice President of the Southern Economics Association, and is associated with the American Enterprise Institute, Independent Institute, and the Technology Policy Institute. Rubin has been Senior Staff Economist at President Reagan's Council of Economic Advisers, Chief Economist at the U.S. Consumer Product Safety Commission, Director of Advertising Economics at the Federal Trade Commission, and Vice President of Glassman-Oliver Economic Consultants, Inc. He has taught economics at the University of Georgia, City University of New York, and George Washington University Law School. Rubin has written or edited 11 books and has published more than 150 articles and chapters on economics, law, regulation, and evolution in leading journals, and he also frequently contributes to the *Wall Street Journal* and other national newspapers.

EDITORIAL INTRODUCTION

TASERS IN THE MEDIA

Conducted energy devices and criminal justice policy

Steven Chermak, *Senior Editor*

Michigan State University

S amuel Walker's (1993) book, *Taming the system: The control of discretion in criminal justice, 1950–1990*, provides time-tested insights on the relationship between bureaucratic decision making and criminal justice policy. Published in the early 1990s, the book provides an analysis of existing literature that pertains to the exercise of discretion at specific criminal justice decision points. Walker discussed that often both anticipated and unanticipated consequences develop after new policy changes are implemented—a conclusion that is consistent with many other classic and contemporary work that points to the instabilities of policy reform. Legislatures, bureaucratic organizations, and courts often initiate policy change to control discretion, and sometimes these efforts are successful. For example, Walker (1993: 147) referred to the control of deadly force in policing as having "enormous significance," and that the adopted rules reduced the overall number of police shootings and helped minimize racial disparities in the use of deadly force. These changes, in addition, had few unanticipated consequences because officers killed or injured in the line of duty did not increase after the adoption of new deadly force policies. New policies can, thus, positively initiate change in criminal justice organizations, and documenting the "law on the books" provides a good starting point for understanding better the impacts of policy change. Such changes, however, also can result in a bureaucratic chain reaction that subverts, undermines, or delegitimizes the intent of the new rule or decision. Personnel will attempt to break or bend any new rule so that the impact is consistent with their accepted and expected routines.

The conflict between "law on the books" and "law in action" provides one element of the policy context for the important debate engaged in this section of the journal. The use of conducted energy devices (CEDs), the most popular of which is the TASER (Taser International, Inc., Scottsdale, AZ), is widespread within the law-enforcement community—nearly 12,000 law-enforcement agencies have acquired the devices (National Institute of Justice [NIJ], 2008). It could be argued that the adoption of such devices (and other less lethal weapons) is a sensible policy change that was adopted to manage better one of law enforcement's most

Direct correspondence to Stephen Chermak, School of Criminal Justice, Michigan State University, 560 Baker Hall, East Lansing, MI 48824 (e-mail: chermak@msu.edu).

volatile situations—the application of force. Some amount of force must be used for every arrest made by a law-enforcement officer, but the level required varies greatly, and its application potentially could escalate quickly depending on characteristics of the situation, suspect, or the officer. Research indicates that the arrest decision is potentially dangerous to law-enforcement officers and suspects. According to the National Institute of Justices' (2008) recent overview, it was concluded that 1,095 arrest-related deaths occurred between 2003 and 2005 while using force. It is clear that new policy, training, and equipment could be of considerable value in reducing the risks related to officer arrest and force decisions. The use of the TASER has great potential, but a key issue in regard to its application is whether suspects are significantly more likely to be injured or killed by law enforcement.

New research seems to support the conclusion that, like the administrative rule-making changes that positively impacted the exercise of deadly force, the use of TASERs has been a positive policy change that produced few unanticipated consequences. For example, the NIJ study (2008) was directed by a panel that represented many disciplines, which included medical, criminal justice, and sociology. This distinguished panel of experts conducted mortality reviews of CED-related deaths, which included an "analysis of autopsy and toxicology results, findings from the scene investigation, post-exposure symptomatology, post-event medical care, and the extent of nature disease presented" (2008: 2), systematically reviewed extant medical knowledge, and consulted other experts. Although the group notes that open questions still need to be empirically examined, it concludes (2008: 3) the following:

- No evidence of a high-risk of injury or death exists from the use of these devices.
- The chances of serious injury or death are low.
- The devices may produce secondary effects that result in death.
- Medical evidence does not support the notion that the devices pose a risk for cardiac arrest.
- Subjects can breathe while exposed to these devices.

Although evidence supports the conclusion that the adoption of TASERs generally can be considered good policy, considerable political and public discussion has raised significant concerns about the use of these devices. Amnesty International (2007), for example, reported that nearly 300 people have died from TASERs in the United States and Canada since 2001. Although these deaths have occurred, the TASER does not seem to be the primary cause in most cases. However, because of several high-profile cases, the media have been interested in both fatal and nonfatal uses of the TASER. Media coverage of cases in which the TASER was deployed outside policy and caused death or serious injury seem to define the public policy discussion about its use. Media reports criticize the use of TASERs, highlight its problems, note other unrepresentative cases, and show public protests that result from an incident. The presentation of these issues establishes another key aspect of the policy context surrounding this debate: how the media intersects policy debates and can force policymakers to adopt rules or specific rule changes. The media reports on the controversial use of these weapons, which emphasizes

the most contentious elements of the debate. Because the media interest in this topic has been significant, it is likely that many or most TASER-related cases currently receive at least some coverage in the news media. This coverage provided the opportunity for Michael D. White and Justin Ready (2009, this issue) to do an innovative study to help us better understand how the "law works in action" in regard to the deployment of TASERs by law-enforcement officers.

White and Ready's (2009) study compares fatal with nonfatal uses of the TASER by law enforcement. The study is valuable for several reasons. First, little academic work has been done on this topic because access to data is limited. White and Ready searched newspaper archives for media reports of TASER incidents, identified several cases, and analyzed the documents to obtain suspect and incident characteristics. The study shows that the innovative application of open source information can be valuable, especially when access to other data sources is limited. Two policy essays that were written in response to this article, however, discuss the limitations of using media data studies. Gregory B. Morrison (2009, this issue) discusses how the newsworthiness of an incident ultimately determines the type of information available for analysis. Media organizations have limited space and time when summarizing the events that occur in a day, and source organizations might strategically limit access to certain incidents and accounts. Although it would be difficult for a police organization to "cover up" an in-custody death, police organizations would attempt to control the nature of the public discussion about the case. Moreover, as Morrison emphasizes, it is the bizarre, unrepresentative case that is not only likely to be covered but also emphasized in the news. Robert J. Kaminski (2009, this issue) provides an excellent summary of research on this topic organized by study methodology and provides a good reminder that the conclusions from this study must be couched in the larger body of research. The important conclusion is that the study of TASER and police use of force remains an important and growing area of research, and approaches vary to study this topic. In addition, Kaminski recommends case control as a potential approach for future study of this topic.

This study as well as the comments in the policy essays are valuable for the consideration of policy implications, but it is important to consider these policy implications along with the limits of the study. Robert J. Bunker's (2009, this issue) essay provides a critical reminder that a disconnect often occurs between research and the street. His essay speaks to the reality that the use of the TASER is the result of a dynamic decision-making process—situations that escalate and can change quickly with little prior information about the suspect. It, thus, makes sense to pursue this line of research aggressively and examine more closely the strengths and limitations of using the media as a primary data source for this type of research (perhaps starting with a list of incidents from an official source and then examining which incidents were covered and which were not) as well as expanding the number of studies that have used other types of data (case control, for example) and examining the use of TASERs in conjunction with, or instead of, other types of less lethal weapons. With more research, and thorough consideration about the realities of police street decision making, great potential abounds for policy to inform practice.

References

Amnesty International. 2007. *Amnesty International's concerns about TASER use: Statement to the U.S. Justice Department inquiry into deaths in custody*. London, UK: Amnesty International.

Bunker, Robert J. 2009. Should police departments develop specific training and policies governing use of multiple TASER shocks against individuals who may be in vulnerable physiological and psychological states? *Criminology & Public Policy*. This issue.

Kaminski, Robert J. 2009. Research on conducted energy devices: Findings, methods, and a possible alternative. *Criminology & Public Policy*. This issue.

Morrison, Gregory B. 2009. Conducted energy weapons: Learning from operational discretion and encounter outcomes. *Criminology & Public Policy*. This issue.

National Institute of Justice. 2008. *Study of deaths following electro muscular disruption: Interim report*. Washington, DC: Office of Justice Programs.

Walker, Samuel. 1993. *Taming the system: The control of discretion in criminal justice, 1950–1990*. New York: Oxford University Press.

White, Michael D. and Justin Ready. 2009. Examining fatal and nonfatal incidents involving the TASER: Identifying predictors of suspect death reported in the media. *Criminology & Public Policy*. This issue.

Steven Chermak is a professor of criminal justice at Michigan State University. His research interests include identifying effective strategies for reducing crime and violence, policing, domestic terrorism, and media coverage of crime and justice. His current research projects include a National Institute of Justice-funded project that examines the impact of a violence-reduction strategy on gun violence in Pittsburgh, a study that is building a national database of criminal incidents that involve far-right extremists, and another project building a terrorism data archive (www.icpsr.umich.edu/TPDRC). Chermak is the author of *Victims in the news: Crime and the American news media* (Westview Press, 1995), in which he examines how the news production process affects the ways in which crime, victims, and criminal justice are presented to the public. His second book, *Searching for a demon: The media construction of the militia movement* (Northeastern University Press, 2002), explores media coverage of the militia movement after the Oklahoma City bombing.

Examining fatal and nonfatal incidents involving the TASER

Identifying predictors of suspect death reported in the media

Michael D. White
Justin Ready
Arizona State University

Research Summary

According to TASER International, nearly 10,000 police departments in the United States have deployed the TASER as a less lethal force alternative in some capacity. Despite the TASER's increasing popularity, serious questions have been raised about the device's physiological side effects; in particular, Amnesty International has reported that more than 300 people have died after being subjected to the TASER. Although a growing body of research has examined the physiological effects of the TASER on animals and healthy human volunteers in laboratory settings, there has been virtually no empirical analysis of "real-world" fatal and nonfatal TASER cases simultaneously. This article examines all media reports of TASER incidents from 2002 to 2006 through a comprehensive review of LexisNexis and New York Times archives. We compare TASER incidents in which a fatality occurred to TASER incidents in which a fatality did not occur and then employ multivariate analyses to identify the incident and suspect characteristics that are predictive of articles describing TASER-proximate deaths.

Policy Implications

Several suspect factors were significantly associated with the reporting of a fatal TASER incident, including drug use (but not alcohol), mental illness, and continued resistance. Multiple deployments of the TASER against a suspect was also associated with the likelihood of the article describing a fatality—especially if the suspect was emotionally disturbed—

which raises the possibility that the risk of multiple shocks might not be uniform for all suspects. More research is needed to explore the relationship between mental illness, drug use (illicit or therapeutic), continued resistance, and increased risk of death. In the meantime, police departments should develop specific policies and training governing the use of multiple TASER shocks against individuals who could be in these vulnerable physiological and psychological states.

Keywords

TASER, conducted energy device, police use of force, less lethal weapons, media

In testimony before the U.S. Department of Justice inquiry into "deaths in custody following electro-muscular disruption," representatives from Amnesty International (2007: 4) stated:

> AI has collected data on more than 290 cases of individuals in the United States and Canada who since 2001 have died after being struck by police Tasers. While in most cases medical examiners have attributed death to other factors…AI has identified at least 20 cases where coroners have found the TASER served a causal or contributory factor in the death and other cases where the TASER was cited as a possible factor in autopsy reports.

In July 2008, a grand jury in Louisiana indicted a police officer on a manslaughter charge involving the death of a suspect who was shocked nine times with a TASER. During the grand jury proceeding, the District Attorney said, "It is our intention to show at trial that Mr. Nugent [the officer] caused the death of Baron Pikes by Tasing him multiple times, unnecessarily and in violation of Louisiana law" ("Officer charged in death of Tasered man," 2008). Alternatively, at the conclusion of the Department of Justice inquiry on deaths after electromuscular disruption, the National Institute of Justice published a special report stating that "there is no conclusive medical evidence within the state of current research that indicates a high risk of serious injury or death from the direct effects of CED [TASER] exposure" (National Institute of Justice, 2008: 3).[1]

These recent events illustrate the ongoing controversy surrounding the physiological effect of the TASER. Although available evidence suggests that deaths after TASER deployment (i.e., TASER-proximate deaths) are rare, there is little understanding of why fatalities occur in a small

---

1.   Conducted energy device (CED) is the general term used to describe the TASER and other versions of the device. Even though TASER dominates the CED market with approximately 95% of sales in the United States, competitors include Stinger Systems and Law Enforcement Associates. Stinger Systems has sold 12,000 weapons since 2000. Law Enforcement Associates introduced their CED in March 2005.

number of cases and how—if at all—the TASER contributed to those deaths.[2] Unfortunately, researchers' ability to explore causes of TASER-proximate deaths has been constrained by two data limitations. First, this emerging technology is not used in a vacuum and laboratory studies on the physiological effects of the device cannot control for circumstances that develop in a natural setting, such as the presence of drugs and alcohol, poor health of suspects, physical confrontations, the use of handcuffs and other less lethal weapons, and elevated levels of adrenaline and physical exertion. Amnesty International (2007: 4) recognized the contributing role of these real-world circumstances, explaining "most of the individuals who died were agitated, disturbed and under the influence of drugs, and/or had underlying health problems such as heart disease."[3] The degree to which the TASER—when combined with one or more of these circumstances—increases the risk of death remains unknown.

Second, empirical research to date has failed to study fatal and nonfatal TASER incidents simultaneously. Although studies involving coroners' reports examined only those cases in which suspects died, laboratory research and studies using police use-of-force data focused on deployments in which most suspects survived. Neither approach can produce variation in the outcome that is required to develop predictive models with advanced statistical methods—that is, to identify predictors of suspect death. As a result of these two limitations, critical questions about fatal and nonfatal incidents remain unanswered. For example: Do TASER incidents that result in death differ in notable ways from nonfatal incidents? If so, what are the distinguishing characteristics of TASER deployments that result in death? And how do specific risk factors alone—or in combination with other circumstances—increase the likelihood of suspect death? These questions are cause for concern in communities and police departments across the United States because the TASER has replaced chemical sprays and impact weapons as the preferred less lethal alternative for U.S. police.

This article explores these questions by analyzing data collected from all media reports describing police use of the TASER from 2002 to 2006 (N = 521), which were generated through a broad search of LexisNexis and the *New York Times*. We used media reports because they are the only data source that provides information about the circumstances surrounding fatal and nonfatal TASER deployments. First, we used bivariate analyses to compare news reports that described fatal deployments (n = 188) with articles describing nonfatal deployments (n = 333). Then we used multivariate analyses, including logistic regression and CHAID (a form of segmentation modeling) to identify suspect and incident characteristics that are significant predictors of news reports involving TASER deaths. The article concludes with a discussion

2. "TASER-proximate" death is the term the authors use to describe an incident where the suspect was struck with a TASER and subsequently died, but no definitive evidence exists to identify the TASER as the cause of death.

3. In the Louisiana case cited above, the coroner's report found that Pikes was handcuffed and lying on the ground when the police deployed the TASER. He was also physically exhausted after running from the police and had a history of cocaine involvement.

of implications for police policy, training, and practice, as well as the ongoing debate about police use of the TASER.

Prior Research

The Context: Police and the Use of Force

Police officers have legal authority to use force in a wide range of situations, including empty hand or physical force, less lethal weapons (i.e., baton, pepper spray, or CED) and, as a last resort, the use of a firearm (Walker and Katz, 2002). Despite its central role in policing (e.g., Bittner, 1970), research indicates that police use of force is statistically rare, occurring in approximately 1% of police–citizen encounters (Bureau of Justice Statistics, 1999).[4] That said, because there are approximately 43 million police–citizen encounters in a given year, an estimated 421,000 use-of-force incidents occur annually, which translates into approximately 1,100 incidents in a typical day.

The application of physical force may potentially have devastating consequences, not only for the suspect and the officer but also for the law-enforcement agency, the community, and police–community relations (Fyfe, 1988; Geller and Scott, 1992). For example, the National Advisory Commission on Civil Disorders (1968) concluded that the police were a cause of many riots occurring during the 1960s. More recent examples of police–citizen encounters that triggered social turbulence include the L.A. riots after the acquittal of the four officers who were videotaped beating Rodney King in 1991, the strained community relations in New York City after the shooting of Amadou Diallo in 1999, and the civil unrest after the Sean Bell incident in 2006.

Because of the physical and social harms—as well as the political fallout that may result from abuses of force—police agencies have expanded their alternatives to firearms. The President's Commission on Law Enforcement and the Administration of Justice made this central to the national policing agenda in 1965 when it proposed developing a range of less lethal weapons. Throughout several decades, advances in technology have led to the development of force alternatives such as oleoresin capsicum (OC; i.e., pepper) spray, impact weapons, foams, ballistic rounds, nets and, most recently, conducted energy devices (CEDs; Wrobeleski and Hess, 2003). This expanding arsenal is intended to provide officers with a broader range of options when a situation requires the application of force but has not escalated to the point when deadly force is necessary, which therefore reduces the likelihood of serious injury and death to officers and suspects.

During the 1990s, OC spray became popular among police agencies in the United States, and this trend serves as the backdrop for the current work on CEDs because many of the same issues and concerns have been raised (Smith and Alpert, 2000). Specifically, controversies surrounding police use of OC spray have centered upon its use on passive resisters, disproportionate

---

4. This estimate becomes much greater when handcuffing and verbal commands are included as use of force.

use on racial minorities, and potential health risks (Kaminski, Edwards, and Johnson, 1999). Several studies have examined the effectiveness of OC spray and concluded that the weapon effectively incapacitates the suspect in most deployments, results in fewer officer injuries, and reduces police reliance on other types of force (Gauvin, 1994; Kaminski, Edwards, and Johnson, 1998; Lumb and Friday, 1997; Nowicki, 1993).

The TASER

CEDs are becoming a visible component of U.S. policing and media depictions of police behavior. The most popular CED is the TASER (manufactured by TASER International, Scottsdale, AZ), and recent estimates indicate that more than 10,000 law-enforcement agencies have purchased and deployed TASERs in some capacity (most commonly, the M26 and X26 models). The TASER fires two probes at a rate of 180 feet per second, and when striking the subject, delivers a 50,000-volt shock during a 5-second cycle (Vilke and Chan, 2007). Although the technology has the capacity to cause acute discomfort, it is not designed to be a pain compliance weapon:

> CEDs work by incapacitating volitional control of the body. These weapons create intense involuntary contractions of skeletal muscle, causing subjects to lose the ability to directly control the actions of their voluntary muscles. CEDs directly stimulate motor nerve and muscle tissue, overriding the central nervous system control and causing incapacitation regardless.… This effect terminates as soon as the electrical discharge is halted. Immediately after the TASER shock, subjects are usually able to perform at their physical baseline (Vilke and Chan, 2007: 349).

Researchers have not kept pace with the rapid growth of this technology and, until recently, most of the available information about the TASER came from the CED industry, internal police reports, and publications from civil and human rights organizations such as Amnesty International and the American Civil Liberties Union. As adoption of the technology has spread, however, questions have emerged in three general areas: (1) policy (When and against whom is it appropriate to use the TASER?), (2) effectiveness (Does the device work as intended?), and (3) physiological impact (Does the TASER increase the risk of injury or death?). These issues are discussed below, but greater emphasis is placed on physiological effect because it is the focus of this article.

When is it appropriate to use the TASER? Police agencies provide guidance to officers on the appropriate use of force through a force continuum that describes the verbal and physical actions a police officer can take in response to different levels of suspect resistance and behavior. Police departments have varied considerably in terms of where they place the TASER on the force continuum, particularly in the earlier years when the device first became available. A central issue is whether the TASER should be used on suspects who are not following verbal commands or who are passively resisting police efforts. This issue became highly controversial when the media captured images of two separate incidents in which the police used the TASER on a

college student who was being passively resistant. The first incident occurred in a University of California at Los Angeles library in November 2006 and the second took place during a press conference for Senator John Kerry at the University of Florida in 2007. Other questionable police practices spotlighted in the media include use of the device on vulnerable populations, such as the physically and mentally disabled, pregnant women, children, and the elderly. The Police Executive Research Forum (PERF, 2005) and the International Association of Chiefs of Police (IACP, 2005) have issued policy recommendations to offer guidance to agencies in their use of CEDs. Both the IACP and PERF suggest that CEDs be used against only those who are actively resisting, that they not be used against minors or the elderly except in emergency situations, and that each deployment be closely supervised.

Does the TASER work effectively? Research examining the effectiveness of the device in the field has focused on two questions: (1) Does the TASER have the intended physiological effect, thereby terminating suspect resistance; and more generally, (2) does use of the device reduce suspect and officer injuries? Unfortunately, limited empirical research is available to answer these questions. With regard to suspect resistance, field data analyzed by TASER International (2006) and internal evaluations by police agencies (Seattle Police Department, 2004) place the effectiveness rate of the TASER somewhere between 80% and 94%. White and Ready (2007) analyzed TASER deployment records from the New York City Police Department and found that suspects stopped resisting and were successfully incapacitated in 86% of incidents occurring during a 4-year period. In a follow-up study, White and Ready (2009) found that the effect of the device on suspect resistance was mitigated by several factors, including suspect weight, intoxication, and the distance between the suspect and the officer.

Several police agencies that have implemented CEDs on a broad scale have later reported reductions in injuries sustained during police–citizen contacts. Police departments in Austin, Texas; Putnam County, Florida; and Cincinnati, Ohio, reported reductions in injuries to suspects and officers after adopting the TASER (Jenkinson, Neeson, and Bleetman, 2006; TASER International, 2006). Smith, Kaminski, Rojek, Alpert, and Mathis (2007) examined injuries to suspects and officers in two departments and concluded that use of CEDs was associated with reduced odds of injury in one department but not the other. The authors noted that "not every agency's experience will be the same regarding CED use and injuries" (Smith et al., 2007: 439).

Does the TASER increase the likelihood of suspect death? Amnesty International has spearheaded efforts to raise awareness of the potentially harmful physiological effects of the TASER. Concern about the possible health risks associated with CEDs has prompted a growing body of research involving reviews of coroner reports in death cases, comprehensive reviews of empirical research, and biomedical research using animals and healthy human volunteers. Kornblum and Reddy (1991) examined 16 deaths after TASER deployments and determined

that drug overdose was the cause of death in most incidents.[5] Ordog, Wasserberger, Schlater, and Balasubramanium (1987) examined 218 emergency room cases that involved suspects who were shocked with a TASER; three suspects died and all had PCP in their systems. Strote, Campbell, Pease, Hamman, and Hutson (2006) examined autopsy reports in 28 TASER-related deaths and found that the device was not identified as the direct cause of death in any of the cases; however, the authors noted that it was listed as a potential contributing factor in 21% of the fatalities (see also Canadian Police Research Centre, 2005; Joint Non-Lethal Weapons Human Effects Center of Excellence, 2005).

Several studies have examined the physiological impact of the TASER on animals (e.g., pigs and dogs) and healthy human volunteers, focusing specifically on cardiac rhythm disturbances such as ventricular fibrillation (e.g., McDaniel, Stratbucker, and Smith, 2000; McDaniel, Stratbucker, Nerheim, and Brewer, 2005; Roy and Podgorski, 1989; Stratbucker, Roeder, and Nerheim, 2003). For instance, both Levine, Sloane, Chan, Vilke, and Dunford (2005) and Ho et al. (2006) monitored heart function in human volunteers who were subjected to the TASER and found no evidence of changes in heart rhythm or functioning. In their review of this research, Vilke and Chan (2007: 353) concluded the following:

> The potential for life-threatening cardiac dysrhythmias or cardiac muscle damage to occur as a result of the electrical discharge from current TASER devices appears to be low based on the available studies. However, there may be theoretical risks to patients with pacemakers or underlying cardiac disease, and the effect of recurrent or prolonged TASER discharges remains unclear.

Vilke and Chan (2007: 353) noted that existing research has yet to investigate "non-cardiac effects" of the TASER, including the device's effect on metabolism (i.e., potassium, sodium, and pH levels in the blood) and respiration (i.e., carbon dioxide elimination and respiratory rate).

Laboratory studies using animals and healthy human volunteers represent an important step in understanding the physiological effect of the TASER, but the implications of this research may be limited by fundamental differences between the test subjects and those individuals most likely to experience the TASER in a real-world setting. That is, suspects who are typically shocked by the TASER during police–citizen encounters are less likely to be healthy, they are not voluntary participants, they are frequently under the influence of drugs or alcohol, they often have preexisting physical or mental conditions, and many are physically resistant during the TASER deployment. Any of these conditions occurring in a natural setting can lead to an increased risk of physiological harm. In October 2007, Wake Forest University Medical Center released their findings from a large-scale evaluation that involved physician reviews of 1,000 real-world TASER deployments, thereby avoiding many limitations associated with laboratory research using animals and human volunteers. The Wake Forest research team coordinated with

5.   In one of the cases, the authors concluded that the TASER could have contributed to the suspect's death because he had a history of cardiac disease, although the suspect also had lethal levels of PCP in his system (see also Vilke and Chan, 2007).

six law-enforcement agencies. When an officer in one of those agencies deployed the TASER, the suspect was transported to a hospital for evaluation by a research team doctor. The study found that 99.7% of suspects had minor or no injuries.[6] William Bozeman, the study's lead investigator, reported:

> The injury rate is low and most injuries appear to be minor. These results support the safety of the devices (Wake Forest University Baptist Medical Center, 2007).

Summary

Although police frequently rely on the TASER when they experience resistance from crime suspects and disorderly individuals, many questions are unanswered concerning the physiological effects of the device when deployed in a natural setting. Given the absence of a causal link to suspect fatalities, why do deaths occur in a small number of incidents? Also, are there identifiable risk factors involving the behavior of the suspect or the circumstances of the encounter that increase the likelihood of death? If so, what are the implications of those risk factors for police policy and training? The following analysis begins to explore these issues using national media data describing fatal and nonfatal TASER incidents.

Methodology

We conducted a broad search of media reports via LexisNexis and *New York Times Select* using keyword searches to identify all articles involving the TASER from January 2002 through December 2006.[7] We relied on media data because such data are the only viable data source with detailed information on fatal and nonfatal cases. After the universe of news reports was identified ($N = 691$), we recorded information for 68 variables relating to the content of the articles and the circumstances under which police used the TASER, placing special emphasis on characteristics of the suspect, officer, and events leading to the deployment.[8] Several categories of cases were excluded from the analysis. First, our study focused only on cases where a police officer deployed the TASER. Two types of news stories did not satisfy this criterion: business reports for TASER International (i.e., stock reports) and articles describing general trends or research relating to police use of the TASER. Additionally, a small number of news reports (approximately 30; 5.4%) were eliminated because of missing information. These stories were typically one or two sentences long and provided little detail about the incident. The authors set aside these reports because their inclusion added little value to the descriptive or multivariate

---

6. Two suspects died, but neither death occurred as a result of the TASER.

7. We used a range of keywords to conduct the database searches. These terms included "TASER," "CED," and "electronic stun device."

8. We developed our coding instrument based on a TASER deployment reporting form used by a major metropolitan police department. Also, the instrument was focused on types of information typically included in articles about the TASER. Unfortunately, certain types of information such as the race, height, and weight of the suspects; the distance to the suspect; and many officer characteristics could not be included in the final analysis because a large proportion of the news reports did not contain this information.

analyses that are the central focus of the article. Finally, duplicate cases were excluded to prevent certain incidents from being overrepresented in the analysis. Duplicate cases were defined as any news report that contained the same information as another report that described a specific incident in which police used the TASER on a suspect (often published on the same date).[9] In all, 170 articles were excluded from the analysis.

The content analysis was based on 521 nonduplicate articles, each of which described an incident in which police used the TASER on a suspect. The dependent variable in this study was whether the article described a TASER incident resulting in death—yes or no. In all, 36% of the articles (n = 188) described an incident in which the suspect died after being shocked with a TASER and the remaining 64% (n = 333) of the articles describe nonfatal TASER incidents. Articles that described TASER-proximate deaths were overrepresented in the data, as research indicates that deaths after TASER use are rare (Wake Forest University Baptist Medical Center, 2007).

Limitations and Considerations of the Data Source

Three types of potential bias are associated with use of media data on TASER events, and each warrants some discussion. The first potential bias stems from differences in the "newsworthiness" of fatal and nonfatal incidents involving the TASER. That is, nonfatal incidents are less likely to be reported in the news because they lack the inherent "drama" associated with fatal incidents. Although the extent of this bias remains unknown, two issues might mitigate its impact in this study. First, given the high level of interest and debate surrounding the device, it could be that nearly all TASER incidents—even nonfatal cases—have the same likelihood of being reported in the media and therefore being captured in our data.[10] Second, media reporting of nonfatal incidents could be most likely when some other interesting characteristic is associated with the event, such as repeated use of the device against a suspect or use against vulnerable populations (e.g., the mentally ill and elderly). In effect, our data-collection approach likely captured similar types of TASER events—those with newsworthy characteristics—that had different outcomes—fatality or no fatality. In short, these nonfatal-though-newsworthy events may actually represent the best comparison for fatal TASER incidents.

The second potential bias concerns how the measures were operationalized and the quality of information recorded from the media reports. We recognized this limitation from the outset and attempted to address it in several ways. First, police reports and interviews of police officers are the primary source of information used by news reporters for most crime and justice issues, and use of the TASER is no exception. In this sense, concerns regarding quality in media data mirror concerns with official police data and other data sources. Second, we examined the extent of missing data across fatal and nonfatal articles and found notable consistencies, especially

9. The duplicate cases were identified by cross-referencing the articles using the date of the incident, city, suspect's name, and the newspaper in which the article was published.
10. Although the extent of news coverage may vary based on the outcome (fatal or nonfatal), our data collection approach captures all those media-reported incidents.

among some of the key variables. For example, the number of times the TASER was deployed during the incident was reported in nearly all news stories (i.e., this information was missing in 4% of nonfatal incidents and 3% of fatal incidents). Alternatively, officer and suspect race were almost never reported, regardless of incident outcome (i.e., in 1% in both types of articles). Articles that described fatal and nonfatal incidents were also similar in their reporting of suspect mental illness (missing in 22% and 28%, respectively) and drug use (missing in 27% and 33%, respectively). At a minimum, similarities in missing data suggest some degree of consistency in data quality across article type (fatal and nonfatal incidents).

Last, each question in the content analysis instrument was designed so that the coder recorded whether a specific piece of information was reported in the article. The coders were not allowed to speculate about any characteristics of the police-citizen encounters.[11] Moreover, in cases in which multiple articles were published about the same event, we compared stories to corroborate "facts" and identify discrepancies.[12] As a result, the main source of error was missing information that could result in a false negative—such as when a coder correctly indicated that an article did not report that the suspect was mentally ill when in fact he was. We attempted to minimize this type of error by focusing the data collection and analyses on items that were expected to yield reasonably accurate data.[13]

The third potential concern is source bias; that is, media representations of police use of the TASER could differ from actual TASER use. Some circumstances could mitigate concern over this bias. First, currently little research is available on police use of the TASER, particularly studies comparing fatal and nonfatal incidents (although the NIJ funded a six-site study in 2008–2009 and other studies are likely underway). Second, although an extensive body of literature has documented media bias in depicting police practices (Chermak, 1995; Surette, 1998), research also suggests that the police are much less effective in "controlling the media message" when the use of force is involved (Chermak, McGarrell, and Gruenewald, 2006; Tuch and Weitzer, 1997). Third, Ready, White, and Fisher (2008: 27) compared news reports with police records on TASER deployments and found "noteworthy consistencies across data sources with regard to many suspect and incident-related characteristics."

In short, this study reflects an effort to identify predictors of TASER-related deaths as reported in the news media. Although the actual influences of the biases described above is

---

11. The content analysis was completed by us and six graduate students from the John Jay College of Criminal Justice. For example, one item in the instrument asks whether the article described the suspect as mentally ill or emotionally disturbed (yes/no). The question did not measure whether it seemed that the suspect was mentally ill based on the narrative describing the encounter. The researchers were instructed to only check "yes" if the article explicitly described the subject as possessing this trait. Thus, the unit of analysis is the article itself, rather than the encounter.

12. In cases where two articles provided inconsistency or contradictory information (which was infrequent), the authors coded information from the article that provided the most detail.

13. The authors also compared these items with available data on the TASER from one police department as a measure of convergent validity.

unknown, we believe that the article makes a substantive contribution to this underdeveloped area of research.[14]

Analysis

We first compared articles describing fatal and nonfatal TASER incidents on the bivariate level across various suspect and incident characteristics. We then used two multivariate approaches—logistic regression and CHAID (a form of segmentation modeling) to identify significant predictors of whether an article describes a TASER-related death. Binary logistic regression was employed because the dependent variable is dichotomous with a "yes" or "no" response. Similar to logistic regression, CHAID predicts the probability of an event occurring, but the method relies on different assumptions and properties and uses segmentation modeling to achieve the task. CHAID divides a population into increasingly homogenous segments that differ based on the dependent variable—in this case, whether the article describes a TASER-proximate death (Jones, Harris, Fader, and Grubstein 2001). The resulting segments are mutually exclusive and exhaustive. As the analysis proceeds, the best predictor is selected among a particular subgroup of cases based on chi-square analyses.

CHAID analysis was employed in this study because it offers several practical advantages. First, CHAID is useful for identifying different sets of predictors for subsets of a population (Jones et al., 2001). For example, predictors of a fatal outcome may be different for intoxicated and sober suspects, and CHAID facilitates the identification and exploration of these interactions. Second, the use of multiple methods allows researchers to "triangulate" their findings or to identify inconsistencies across techniques. Third, CHAID has been used in the study of police practices previously, including the use of force and arrest decisions (Kane, 1999; White, 2002, 2006, 2008). Finally, an added benefit of CHAID is the user-friendly visual representation of complex analyses and relationships for practitioners and policymakers.

Findings

Table 1 provides a frequency distribution for the basic content of all LexisNexis and *New York Times* articles about the TASER from January 2002 through December 2006. LexisNexis articles account for 84.8% (*n* = 586) of the cases, whereas the remaining 15.2% (*n* = 105) of the cases were derived from the *New York Times*. Overall, 75.4% of the news reports describe a specific incident in which the police used the TASER on a suspect. The volume of news publications about the TASER increased from 24 (3.5%) in 2002 to 179 (25.9%) in 2004, and then peaked at 338 articles (48.9%) in 2006. The regions of the country that were the source of the most articles are the Southeast (*n* = 248) and the West (*n* = 190). The Northeast has produced the

---

14. The authors are grateful to Senior Editor Steve Chermak for assistance in delineating the strengths and weaknesses of the data source.

fewest articles relating to the TASER, accounting for 8.4% (*n* = 58) of cases.[15] As a whole, the news reports offered considerable detail about the circumstances under which officers used the device in the field, the events leading up to the deployment, and the behavior and characteristics of the suspects and officers. However, a relatively small proportion of the articles also discussed when it is appropriate for police officers to use the weapon (17.4%) and the physiological effects (or effectiveness) of the TASER (18.0%).

TABLE 1

Content Analysis of LexisNexis and *New York Times* Articles about the TASER, 2002–2006

Variable	Percent	Valid Cases
Source of the article		
LexisNexis	84.8	586
New York Times	15.2	105
Total	100.0	691
Type of Story		
Incident where police deployed the TASER	75.4	521
Duplicate case or not a TASER deployment*	24.6	170
Total	100.0	691
Did the TASER deployment result in death?		
Yes	36.1	188
No	63.9	333
Total	100.0	521
Did the article discuss when it is appropriate for police to use the TASER?		
Yes	17.4	120
No	82.6	571
Total	100.0	691
Did the article discuss the physiological effects (or effectiveness) of the TASER?		
Yes	18.0	124
No	82.0	567
Total	100.0	691
Year of publication		
2002	3.5	24
2003	6.4	44
2004	25.9	179
2005	15.3	106
2006	48.9	338
Total	100.0	691

15. Many articles published in the *New York Times* were not counted in the Northeast because they duplicated LexisNexis articles that originated in other regions of the country. The states that generated the most news coverage of the TASER are Florida (*n* = 122), California (*n* = 77), and Colorado (*n* = 60; not shown in Table 1). Altogether, these three states accounted for 37.5% of non-duplicate news reports about the TASER during the study period.

Region of country		
Northeast	8.4	58
Southeast	35.9	248
Midwest	18.8	130
West	27.5	190
Not in United States/Not applicable	9.4	65
Total	100.0	691

Comparative Findings

Table 2 provides a comparison of news reports describing fatal and nonfatal TASER deployments during the 5-year study period (2002–2006). The variables in the analysis were separated into two categories: incident and suspect characteristics. Several differences are noteworthy. The articles describing TASER incidents that resulted in death were more likely to involve situations in which the police officer used the weapon more than once on the same suspect (50.8% vs. 23.3% for nonfatal cases), suspect resistance continued after TASER deployment (38.8% vs. 22.7% for nonfatal), the suspect was handcuffed (i.e., in custody) when the TASER was deployed (22.3% vs. 6.1% for nonfatal cases), and the suspect was transported to the hospital (58.6% vs. 29.8% for nonfatal cases).[16] Table 3 shows the relationship between the number of times the TASER is used on the same suspect and a fatal outcome in the news media. Approximately one fourth of the articles that describe incidents where the TASER was used once resulted in suspect death. In the articles where the TASER was used repeatedly on the same suspect, fatal outcomes increased considerably (50–55% for 2 or 3 deployments; 60% for 4 or more deployments).

T A B L E   2

Content Analysis of News Reports Describing Police TASER Use: A Comparison of Fatal and Nonfatal TASER Deployments, 2002–2006

Variables	Nonfatal		Fatal	
	Percentage	Valid Cases	Percentage	Valid Cases
Incident characteristics				
(i.e., circumstances)				
Officer used TASER more than once	23.3%	322	50.8%	181
Suspect continued to resist	22.7%	326	38.8%	178
Other less lethal weapon used before	16.2%	327	22.5%	182
Other less lethal weapon used after	26.1%	326	31.0%	184
Suspect in custody when TASER used	6.1%	330	22.3%	188
Suspect taken to hospital	29.8%	332	58.6%	186
Backup officer(s) present at the scene	71.9%	331	72.2%	180
Suspect characteristics				
Minor or senior citizen	9.6%	332	4.3%	188
Male suspect	91.9%	332	97.9%	188
Intoxicated from alcohol	10.5%	333	8.0%	187

16. Although approximately one third of the articles mentioned that the suspect had attacked the officer, whether or not the officer was injured (and the extent of that injury) rarely was described.

Under influence of drugs	6.3%	332	23.7%	186
Emotionally disturbed or mentally ill	22.9%	332	36.2%	188
Armed with weapon	37.0%	332	16.5%	188
Assaulted an officer	29.7%	330	37.1%	186
Verbal or passive resistance	26.3%	327	22.2%	185

News reports describing fatal and nonfatal TASER incidents also differed in terms of the characteristics of suspects (see Table 2).[17] Suspects in articles that described fatal incidents were more often under the influence of drugs (23.7% vs. 6.3%) and mentally ill or emotionally disturbed (36.2% vs. 22.9%), but they were less likely to be armed with a weapon (16.5% vs. 37.0%). A word of caution is warranted regarding the drug and mental illness variables. Police reports and police officer statements are the primary source of information used by newspaper reporters to document these suspect characteristics. As a result, these variables likely reflect the police officer's assessment of the suspect at the time of the incident based on his or her personal observations and on information collected at the scene (e.g., from family members and witnesses).[18] This information is not based on more definitive tests such as a urinalysis, blood or hair analysis, or clinical assessments, and the reader should bear this in mind. The next section explores whether these incident and suspect characteristics were significant predictors of TASER fatalities in the media when controlling for other factors through multivariate analysis.

TABLE 3

Content Analysis of LexisNexis and *New York Times* Articles, 2002–2006: Number of Times Police Used the TASER by Suspect Death (Number of Valid Cases in Parentheses)

Number of Times Police Used the TASER on a Suspect (i.e., number of cycles)	Suspect Death	
	No	Yes
Once	73.5% (247)	26.5% (89)
Twice	45.4% (44)	54.6% (53)
Three times	50.0% (15)	50.0% (15)
Four or more times	40.0% (16)	60.0% (24)
Total	64.0% (322)	36.0% (181)

17. Nine percent of the articles identified the suspect's race or ethnicity ($n = 51$). Of those 51 articles, 16 involved black suspects, 21 involved Hispanic suspects, and 13 involved white suspects. The failure to report race and ethnicity did not extend to other demographic characteristics, as three quarters of the articles identified the suspect's gender and more than 60% discussed the suspect's age.

18. Interestingly, a large portion of the articles did mention the type of drug the suspect was abusing, typically cocaine or methamphetamine.

Multivariate Findings

Table 4 displays the results from a logistic regression model predicting TASER-related deaths in the news media using suspect and incident circumstances as covariates. The likelihood ratio test for the analysis was significant and the estimate for Nagelkerke *R*-squared indicates that the model accounts for 33.7% of the variation in the outcome variable.

T A B L E   4

**Logistic Regression Model Predicting Police TASER
Deployments Resulting in Suspect Death**

Variables	B	S.E.	Exp(B)
Incident characteristics			
Number of times TASER used	0.018	.038	1.018
Suspect continued to resist after TASER	0.820	.274	2.270**
Other less lethal weapon used before	−0.028	.320	0.972
Other less lethal weapon used after	0.139	.271	1.149
Suspect in custody when TASER used	1.195	.358	3.302**
Suspect taken to hospital	1.217	.250	3.379***
Backup officer(s) present at the scene	−0.429	.272	0.651
Suspect characteristics			
Age of suspect (minor or senior citizen)	−0.491	.484	0.612
Gender of suspect	1.679	.698	5.361*
Intoxicated from alcohol	−0.892	.433	0.410*
Under influence of drugs	1.414	.351	4.112***
Emotionally disturbed or mentally ill	0.562	.261	1.754*
Armed with a weapon	−1.083	.323	0.338***
Level of resistance before TASER (none or verbal)			
Passive physical resistance	0.134	.347	1.143
Active physical	−0.306	.318	0.736
Active potentially lethal	−0.458	.442	0.632
Constant	−2.740	.740	0.065
Log likelihood	477.423		
R^2 (Nagelkerke)	.337		
Chi square	131.794		
d.f.	16		
Significance	.000		
N	470		

Notes. S.E. = standard error.
*p < .05. **p < .01. ***p < .001.

Predictors of articles describing TASER deployments resulting in suspect death include the following:

- Suspect resistance after the weapon was deployed (the likelihood of death was two times greater)
- Suspect was handcuffed (i.e., in custody) when the weapon was deployed (suspect death was more than three times as likely)
- Suspect was transported to the hospital after the TASER was deployed (the likelihood of death was more than three times greater)
- Suspect was under the influence of drugs (suspect death was four times more likely)
- Suspect was emotionally disturbed or mentally ill (suspect death was nearly twice as likely)[19]

Two factors were associated with a *reduced* likelihood of the article describing an incident resulting in death: the suspect being under the influence of alcohol and the suspect being armed with a weapon (in more than half of those articles, the weapon was a knife or other cutting instrument; 23% were armed with a gun, however). Several variables that were not statistically significant are noteworthy, including the number of times the TASER was used on the suspect, the use of other less lethal weapons, the suspect's age, and the level of resistance before the TASER was deployed.
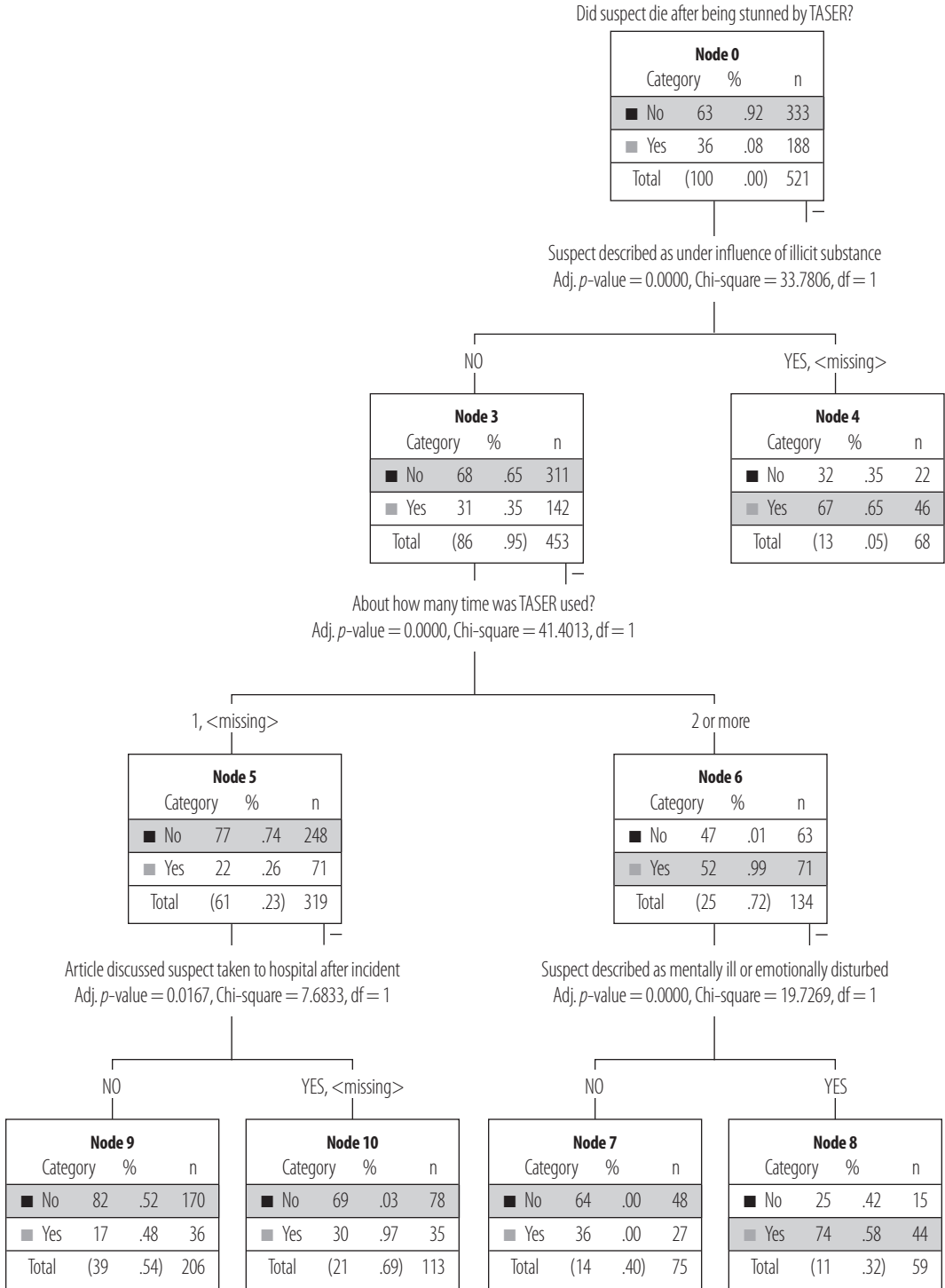
The CHAID findings are displayed in Figure 1. The top cell (or root node) in the CHAID tree includes all 521 articles and highlights the 36.1% in which the suspect died after being shocked with the TASER. The initial split was made based on whether the suspect described in the article was under the influence of drugs (hereafter referred to as "high"; this classification does not include alcohol), thus separating the 521 TASER articles into two cells as follows: those in which the suspect was not high (*n* = 453; 86.9% of the total) and those in which the suspect was high (*n* = 68; 13.1% of the total). The splits in CHAID are made according to differences in the dependent variable: In articles with high suspects, death occurred 67.7% of the time, compared with death occurring 31.4% of the time in articles with suspects who were not high. Another split was made from the not-high cell based on the number of times the TASER was used on a suspect: In articles in which the suspect was not high and the TASER was deployed repeatedly, death occurred 52.9% of the time; compared with suspect death occurring in 22.3% of the articles in which the TASER was deployed once on suspects who were not high.

The next split was made from the cell indicating that the TASER was used more than once. This split is based on whether the suspect was mentally ill or emotionally disturbed: In articles in which the suspect was not high during the encounter but was emotionally disturbed, and in which the TASER was used more than once, suspect fatalities occurred 74.6% of the time, compared with suspect death occurring in 36.0% of the articles in which the suspect was not emotionally disturbed or high (and the TASER was used more than once). The final split is made

19. Suspect gender was statistically significant (*p* > .05), but there were 31 articles describing a female suspect, and 4 involved a TASER-proximate death. We also ran the model without the suspect gender variable and the results were unchanged.

CHAID Analysis with TASER Articles

Did suspect die after being stunned by TASER?

Node 0		
Category	%	n
■ No	63	.92	333
■ Yes	36	.08	188
Total	(100 | .00)	521

Suspect described as under influence of illicit substance
Adj. p-value = 0.0000, Chi-square = 33.7806, df = 1

NO

Node 3		
Category	%	n
■ No	68	.65	311
■ Yes	31	.35	142
Total	(86 | .95)	453

YES, <missing>

Node 4		
Category	%	n
■ No	32	.35	22
■ Yes	67	.65	46
Total	(13 | .05)	68

About how many time was TASER used?
Adj. p-value = 0.0000, Chi-square = 41.4013, df = 1

1, <missing>

Node 5		
Category	%	n
■ No	77	.74	248
■ Yes	22	.26	71
Total	(61 | .23)	319

2 or more

Node 6		
Category	%	n
■ No	47	.01	63
■ Yes	52	.99	71
Total	(25 | .72)	134

Article discussed suspect taken to hospital after incident
Adj. p-value = 0.0167, Chi-square = 7.6833, df = 1

Suspect described as mentally ill or emotionally disturbed
Adj. p-value = 0.0000, Chi-square = 19.7269, df = 1

NO

Node 9		
Category	%	n
■ No	82	.52	170
■ Yes	17	.48	36
Total	(39 | .54)	206

YES, <missing>

Node 10		
Category	%	n
■ No	69	.03	78
■ Yes	30	.97	35
Total	(21 | .69)	113

NO

Node 7		
Category	%	n
■ No	64	.00	48
■ Yes	36	.00	27
Total	(14 | .40)	75

YES

Node 8		
Category	%	n
■ No	25	.42	15
■ Yes	74	.58	44
Total	(11 | .32)	59

from the cell indicating that the TASER was used only once, based on whether the suspect was transported to the hospital or not. In articles in which the suspect was not high and was shocked by the TASER once but was taken to the hospital, suspect death occurred 30.9% of the time, compared with death occurring in 17.5% of the articles in which the suspect was not high, the TASER was deployed only once, and the suspect was not transported to the hospital.

Table 5 provides a summary of the termination cells for the CHAID tree, which features the predictors, cell size, percentage of the total cases, and percentage of the dependent variable. Five termination cells are listed, with the percentage of articles describing suspect death ranging from 17.5% to 74.6%. The CHAID termination cells call attention to a relatively large group of articles (40% of the total) with infrequent suspect deaths (17.5% describing a death), a similar-sized group in which death occurred in one third of the articles (the second and third subgroup combined), and two groups of articles in which suspect death is very common, occurring in 68–75% of the articles. The high-risk groups of articles are characterized by suspects who were under the influence of drugs or emotionally disturbed, and they were subjected to more than one TASER deployment.

TABLE 5

Summary of CHAID Analysis Termination Groups

Article describing a TASER-proximate death indicates…	Number of Articles	% of Total	% Describing Death
1. Suspect not under influence of illicit drug; TASER deployed once; suspect not taken to hospital	206	39.5%	17.5%
2. Suspect not under influence of illicit drug; TASER deployed once; suspect taken to hospital	113	21.7%	31.0%
3. Suspect not under influence of illicit drug; TASER deployed more than once; suspect not emotionally disturbed	75	14.4%	36.0%
4. Suspect under the influence of illicit substance	68	13.1%	67.6%
5. Suspect not under influence of illicit drug; TASER deployed more than once; suspect is emotionally disturbed	59	11.3%	74.6%

Discussion

Notable Nonsignificant Characteristics

It might be useful to explore the implications of the findings by first considering several presumably important characteristics that did not emerge as statistically significant.[20] One controversy surrounding the TASER involves use of the device on vulnerable populations, notably the young and the elderly. Arguably, the physiological makeup of both the young and elderly could elevate the risk of potential negative health effects following TASER deployment. That is, the elderly are more likely to suffer from physical ailments—particularly heart problems—whereas

20. These characteristics are notable either because of prior research on TASER use or because conventional wisdom suggests they might be important.

increased risks for minors often stem from their low body weight and immature physical development. The findings indicate, however, that the suspect's age—measured many different ways—was not predictive of an article describing a TASER-related death. This finding may be a result of the small number of cases that involve those vulnerable populations because only 40 articles described an incident involving a minor or elderly person, and approximately 4% of the TASER-proximate death articles involved such an individual.

Second, some attention has focused on whether the risk of suspect death increases when the TASER is used in conjunction with other less lethal alternatives, such as pepper spray or the baton, often because one of the devices was ineffective in subduing the suspect. The use of other less lethal devices was included as three separate dichotomous variables (run independently), and although nearly 40% of the news reports described the use of another less lethal device, none of the variables was statistically significant. Last, there has been some disagreement about the appropriate placement of the TASER on the force continuum and, more specifically, whether the device should be used in response to passive resistance by a suspect. Approximately one quarter of the news reports described a case that involved a suspect who gave only verbal or passive resistance, but the variable was not predictive of an article describing a fatality.

Significant Predictors

The most important predictor of an article describing a TASER-proximate death was the suspect being the under the influence of drugs. In the logistic regression, when the suspect was high, the article was four times as likely to describe a fatality. In the CHAID model, the "suspect high" variable served as the first split and isolated an end group in which more than two thirds of the articles involved TASER-proximate deaths. This finding is consistent with prior research examining coroners' reports that has documented a link between drug use, TASER deployments, and increased risk of death. Several possible explanations for this relationship are given here. First, despite its recurrent theme in the research, it is conceivable that the relationship between drug use and suspect death is spurious. In many cases, a suspect under the influence of illegal drugs is shocked with the TASER by police and death does not occur. It is possible that some other intervening variable, that often coincides with drug use, increases the risk of death.

Second, the relationship between drug use and increased risk of death could be related to the individual's behavior following the TASER deployment. Continued resistance by the suspect after TASER deployment also emerged as a significant predictor of articles describing fatalities. Suspects who are high might be more likely to continue resisting after the TASER is deployed. This continued physical exertion, which is a result of impaired judgment, could place enough stress on the body that—when combined with the effects of the drug—induces severe physiological responses such as excited delirium. Excited delirium (ED) syndrome is defined as:

> The sudden death of an individual, during or following an episode of excited delirium, in which an autopsy fails to reveal evidence of sufficient trauma or natural disease to explain the death. In virtually all such cases, the episode of excited delirium

[defined as a transient disturbance in consciousness or cognition involving violent behavior] is terminated by a struggle with police or medical personnel, and the use of physical restraint (DiMaio and DiMaio, 2006 :1).

Since 1984, more than 130 cases of fatal, cocaine-associated ED have been reported in the medical and forensic literature, and nearly all have involved use of force (usually by police, although some involve medical staff at psychiatric hospitals; Sztajnkrycer and Baez, 2008). DiMaio and DiMaio (2006: 4) explained that most deaths attributed to ED involve cocaine or methamphetamine abuse and that death is not caused by police use of force; rather, death is likely a product of, "normal physiological reactions of the body to stress gone awry, and to the use of stimulants." In their review of 74 TASER-proximate death cases, DiMaio and DiMaio (2006: 42) concluded that these deaths "almost invariably describe individuals in excited delirium, high on illegal stimulants who die not at the time they are 'Tasered' but sometime after."

A third possible explanation for the drug-use finding involves the combined effect of the illicit drugs and the TASER shock to the suspect's system. Illegal drugs have a wide range of potentially harmful effects on the human body, which include the functioning of major organs such as the brain and heart, the central nervous system, respiration, circulation, and metabolism. These negative effects—which by themselves can be fatal—might become more lethal when combined with the effects of one or more shocks from the TASER. Also, the combined effect of the TASER and an illicit drug may vary by the type of drug used. For example, results from this study suggest that alcohol intoxication has an inverse relationship with the probability of an article describing a TASER-proximate death. Perhaps the risk of death is greater with stimulants, such as methamphetamines and cocaine (by far, the most common type of drugs described in the news reports), than with depressants such as alcohol.

Mental or emotional distress was also a significant predictor in the multivariate models. When the suspect was described as being mentally ill and in crisis, there was a greater probability that the article involved a TASER-proximate death. In particular, the CHAID analysis isolated a small group of articles (n = 59; 10% of the articles examined) in which the suspect was emotionally disturbed or mentally ill and the officers used the TASER more than once. Three quarters of these articles involved a TASER-proximate death. Much like suspects who are under the influence of drugs, those who are mentally ill and in crisis might be less likely to acquiesce to police authority (i.e., continued resistance) because of their emotional state.[21] Continued physical resistance might then prompt the police to apply more cycles from the TASER. The number of times the device is used on a suspect has received considerable attention from civil liberties and human rights organizations because of the potential for excessive force (i.e., Amnesty International indicated that the device could be used to torture a suspect) and because of the risk of injury or death. Vilke and Chan (2007: 353) concluded, "the effect of recurrent

---

21.   Recall that continued resistance by itself was a significant predictor in the logistic regression analysis.

or prolonged taser discharges remains unclear." The findings here raise the possibility that the impact of multiple shocks might not be consistent for all suspects and that the risk of death might be greater for certain subgroups, such as those who are mentally ill. Moreover, suspects who are mentally ill could also have psychotropic drugs in their systems, which increases the risk of death. DiMaio and DiMaio (2006) noted that therapeutic drugs, such as antidepressants and antipsychotics often prescribed to the mentally ill, have been linked to ED.

Several other predictors were associated with a greater likelihood that an article described a fatality. First, if the suspect was transported to the hospital, then there was an increased likelihood that the article described a TASER-proximate death. This finding makes intuitive sense, as suspects who later died would have exhibited symptoms at the scene that led to their transport to the hospital.[22] Second, nearly one quarter of the articles describing a TASER-proximate death involved a suspect in custody. Most of these suspects were handcuffed or being detained in a holding cell. The nature of this relationship is unclear, particularly because in-custody deaths are rare. A recent Bureau of Justice Statistics (2007) study documented a total of 2,002 arrest-related (or in-custody) deaths in the United States from 2003 to 2005. Given that police in the U.S. make approximately 1.3 million arrests each year (Bureau of Justice Statistics, 2005), the number of in-custody deaths during that 3-year period represents less than 0.5% of arrests. Nevertheless, when a suspect dies in police custody, the case generates substantial controversy and often raises questions about excessive force and police brutality. The finding here may simply be a consequence of the "newsworthiness" of in-custody deaths. However, it is worth considering whether use of the device with handcuffed suspects is a health risk in cases in which the suspect's air passages are already restricted (e.g., when the individual is lying on his stomach).

Third, the suspect being armed with a weapon is inversely related to the likelihood that the article describes an incident resulting in death. This finding too may be a consequence of media reporting but in a slightly different way. Although much of the news reporting can be characterized as negative or critical of the TASER, some articles were decidedly more "pro-TASER."[23] In fact, 35 articles describe incidents in which the suspect was armed with a gun (and 79 described suspects armed with knives or other cutting instruments). The emergence of the "suspect armed" variable may be an artifact of this second type of news report, which includes articles that describe "successful" TASER deployments involving armed and dangerous suspects who were thwarted by the less lethal device.

---

22. Some police departments, such as the New York City Police Department, require that all suspects who are shocked with a TASER be transported to the hospital for examination. The recent guidelines by PERF and IACP make similar recommendations.

23. We make this "anti-TASER" assertion based on the content analysis of more than 500 news articles.

Implications for Policing

The implications of this study for police departments center on four issues. First, substantial variation exists across departments with regard to policy, use, and reporting of TASER deployments. Variation in critical issues—such as when the TASER should be used and who it should be used against—has resulted in controversial incidents and could be related to an increased likelihood of fatal outcomes. Although the limited scope of research on the TASER prohibits any sort of discussion related to "best practices," the PERF and the International Association of Chiefs of Police (IACP) offer model policy guidelines involving TASER use. In the absence of "best practices," we believe departments would be well advised to consult these guidelines carefully when crafting their own policies. Adherence to PERF/IACP standards regarding suspect resistance level, use against vulnerable persons, multiple deployments, and reporting practices will likely reduce the potential for controversial and inappropriate deployments.

Second, police departments should routinely collect and analyze data on their officers' TASER use, and they should consider making that data publicly available. With regard to the first point, many police departments require that officers file reports when any degree of force is used, but not all departments do. Kane (2007: 775) noted that the primary motivations for collecting data on use of force is to identify problem officers and he explained that "police departments that collect data on their practices can identify far better than others policies that work, policies that do not work, and areas of organizational behavior that should be regulated." The second point is more contentious—that police should make those data available to the public. Although police are traditionally reluctant to release data, especially sensitive information such as use-of-force data, there are two reasons for public release of TASER data. First, there is a critical need for transparency in this area. Use of force in general, and use of the TASER specifically, can cause long-term damage to a department's relationship with the community. By releasing data to the community, the department can convey a powerful message of accountability and openness. Moreover, Kane argued that police have *an obligation* to share such data because the public funds and "owns" all information generated by the department:

> Currently, however, police departments operate under a paradigm where the burden of proof rests on the members of the public to justify adequately why they want access to police data. The paradigm should change such that the burden of proof is on police departments to justify why the information should be kept unavailable to the public. Until this paradigm shifts, police accountability in the United States will remain stalled, and members of the public will continue to be harmed unnecessarily by the police (Kane, 2007: 779).

Third, police departments should seek partnerships with local researchers to study use and deployment of the TASER. Departments could continue to explore some of the intuitive and counterintuitive findings from this study. In particular, departments could explore the potential differential effects of the TASER with certain types of drugs. Is the device more or less effective on suspects under the influence of co s? And how does physical exertion and active resistance

interact with these substances? Partnerships between police and local researchers could also explore tactical issues such as multiple deployments, the use of multiple less-lethal weapons, and the so-called "velcro effect"—where simply showing the device may produce suspect compliance.[24] Voluntary release of data by police, coupled with analysis by independent researchers, would move knowledge in the field forward and represent a significant step toward transparency with possibly profound implications for police–community relations.

Fourth, police departments should become familiar with the medical research on the effects of the TASER and excited delirium. Officers typically receive training on how to use the TASER in the academy and during in-service "refresher" courses. Information pertaining to medical risks, physiological responses, and high-risk medical situations should be incorporated into these training sessions. In other words, training should move beyond the "tasing" of volunteer police recruits while their fellow classmates look on in amusement—a traditional and still popular method for introducing recruits to the TASER. Clearly, the focus of this training should remain on operational and tactical issues, but trainers should also provide officers with state-of-the-art medical information so that, once on the street, officers are keenly aware of the potential risks and can quickly identify medical emergencies as they occur.

Research and Policy Implications

There are also some important next steps for researchers and funding agencies. First, additional studies should explore the value of media data for studying police field behavior, especially use of force. One must be wary of many biases; nevertheless, research has failed to document the actual extent or pervasiveness of bias. If concerns over these biases can be properly addressed, this methodological approach could then be extended to the study of other types of police field behavior where data are traditionally difficult to obtain (e.g., automobile pursuits and deadly force).

Second, researchers and government agencies would benefit from a continued dialogue on the creation of a national use-of-force reporting system, which would include both lethal and less lethal force incidents (e.g., TASER deployments). Hickman, Piquero, and Garner (2008) made important strides in this area by highlighting the shortcomings of the Police–Public Contact Survey, and policy essays by Klinger (2008) and Smith (2008) offered thoughtful guidance on how best to initiate and structure such a system.[25] Clearly, our understanding of physiological, tactical, and policy issues associated with TASER use would expand immeasurably through analysis of incidents in a national reporting system.

Third, the call for more research in this area coincides with the increased availability of funding through the Economic Recovery Act of 2009. Although the NIJ continues to be a viable source of funding for police research (with a track record in supporting TASER research),

---

24. We thank one of the anonymous reviewers for introducing us to this term.

25. Hickman et al. (2008) also noted that other nations such as New Zealand have already developed national use-of-force reporting systems that the United States could model.

social scientists should seek opportunities in nontraditional venues, such as the National Science Foundation, National Institutes of Health, and the Centers for Disease Control and Prevention. The study of CEDs generally, and the TASER specifically, has been fragmented often with little interaction among the relevant fields. Partnerships among social scientists, physicians (or others in the medical field), pharmacologists, psychiatrists, and police practitioners could broaden this line of research through empirical studies that are both comprehensive and multidisciplinary. Such an approach would be ideal for examining the perceived relationships between mental illness, substance abuse, active resistance, multiple TASER deployments, and an increased risk of death.

References

Amnesty International. 2007. Amnesty International's concerns about TASER use: Statement to the U.S. Justice Department inquiry into deaths in custody. London, UK: Amnesty International.

Bittner, Egon. 1970. *The functions of police in modern society*. Washington, DC: U.S. Government Printing Office.

Bureau of Justice Statistics. 1999. *Contacts between police and the public: Findings from the 1999 national survey*. Washington, DC: U.S. Department of Justice.

Bureau of Justice Statistics. 2005. *Contacts between police and the public: Findings from the 2002 national survey*. Washington, DC: U.S. Department of Justice.

Bureau of Justice Statistics. 2007. *Arrest-related deaths in the United States, 2003–2005*. Washington, DC: U.S. Department of Justice.

Canadian Police Research Centre. 2005. *Review of conducted energy devices*. Ottawa, Ontario: Author.

Chermak, Steven. 1995. Image control: How police affect the presentation of crime news. *American Journal of Police*, XIV: 21–43.

Chermak, Steven, Edmund McGarrell, and Jeff Gruenewald. 2006. Media coverage of police misconduct and attitudes toward police. *Policing: An International Journal of Police Strategies and Management*, 29: 261–281.

DiMaio, Theresa G. and Vincent J.M. DiMaio. 2006. *Excited delirium syndrome: Cause of death and prevention*. New York: Taylor and Francis.

Fyfe, James J. 1988. Police use of deadly force: Research and reform. *Justice Quarterly*, 5: 165–205.

Gauvin, Roy. 1994. Oleoresin capsicum spray: A progress report. Portland, OR: Portland Police Department.

Geller, William and Michael S. Scott. 1992. *Deadly force: What we know*. Washington, DC: Police Executive Research Forum.

Hickman, Matthew J., Alex R. Piquero, and Joel H. Garner. 2008. Toward a national estimate of police use of nonlethal force. Criminology & Public Policy, 7: 563–604.

Ho, Jeffrey D., James R. Miner, Dhamunjaya Lakireddy, Laura L. Bultman, Laura Heegaard, and William G. Heegaard. 2006. Cardiovascular and physiologic effects of conducted electrical weapon discharge in resting adults. Academic Emergency Medicine, 6: 239–243.

International Association of Chiefs of Police (IACP). 2005. *Electro-muscular disruption technology (EMDT): A nine-step strategy for effective deploymen*t. Alexandria, VA: Author.

Jenkinson, Emma, Clara Neeson, and Anthony Bleetman. 2006. The relative risk of police use-of-force options: Evaluating the potential for deployment of electronic weaponry. *Journal of Clinical Forensic Medicine,* 13: 229–241.

Joint Non-Lethal Weapons Human Effects Center of Excellence (HECOE). 2005. *Human effectiveness and risk characterization of the electromuscular incapacitation device—A limited analysis of the TASE*R. Brooks Air Force Base, TX: Air Force Research Laboratory.

Jones, Peter, Phillip Harris, Jamie Fader, and Lorie Grubstein. 2001. Identifying chronic juvenile offenders. *Justice Quarterly*, 18: 479–507.

Kaminksi, Robert J., Steven M. Edwards, and James W. Johnson. 1998. The deterrent effects of oleoresin capsicum on assaults against police: Testing the velcro-effect hypothesis. *Police Quarterly*, 1: 1–20.

Kaminksi, Robert J., Steven M. Edwards, and James W. Johnson. 1999. Assessing the incapacitative effects of pepper spray during resistive encounters with the police. *Policing: An International Journal of Police Strategies and Management*, 22: 7–29.

Kane, Robert J. 1999. Patterns of arrest in domestic violence encounters: Identifying a police decision-making model. *Journal of Criminal Justice,* 27: 65–79.

Kane, Robert J. 2007. Collect and release data on coercive police actions. *Criminology & Public Policy,* 6: 773–780.

Klinger, David A. 2008. On the importance of sound measures of forceful police actions. *Criminology & Public Policy,* 7: 605–617.

Kornblum, Ronald N. and Sara K. Reddy. 1991. Effects of the Taser in fatalities involving police confrontation. *Journal of Forensic Science*, 36: 434–448.

Levine, Saul D., Christian M. Sloane, Theodore C. Chan, Gary M. Vilke, and James V. Dunford. 2005. Cardiac monitoring of subjects exposed to the Taser. *American Emergency Medicine*, 13: S47.

Lumb, Richard C. and Paul C. Friday. 1997. Impact of pepper spray availability on police officer use of force decisions. *Policing: An International Journal of Police Strategies and Management*, 20: 173–185.

McDaniel, Wayne C., Robert A. Stratbucker, Max Nerheim, and James E. Brewer. 2005. Cardiac safety of neuromuscular incapacitating defense devices. *Pacing and Clinical Electrophysiology*, 28: 284–287.

McDaniel, Wayne C., Robert A. Stratbucker, and Rick Smith. 2000. *Surface application of Taser stun guns does not cause ventricular fibrillation in canines*. Proceedings of the Annual International Conference of the IEEE EMBS, Chicago, IL.

National Advisory Commission on Civil Disorder (Kerner Commission). 1968. *Report of the National Advisory Commission on Civil Disorder*. Washington, DC: U.S. Government Printing Office.

National Institute of Justice. 2008. *Study of deaths following electromuscular disruption*: *Interim report.* Washington, DC: Office of Justice Programs, U.S. Department of Justice.

Nowicki, Edward. 1993. Oleoresin capsicum: A non-lethal force alternative. *Law Enforcement Technology*, 20: 24–27.

"Officer charged in death of Tasered man." 2008. *CNN*, August 13. Retrieved December 22, 2008 from cnn.com/2008/CRIME/08/13/la.taser.death/index.htm.

Ordog, Gary, Jonathan Wasserberger, Theodore Schlater, and Subramanium Balasubramanium. 1987. Electronic gun (taser) injuries. *Annals of Emergency Medicine*, 16: 73–78.

Police Executive Research Forum (PERF). 2005. *PERF conducted energy device policy and training guidelines for consideration*. Washington, DC: PERF Center on Force and Accountability.

Ready, Justin, Michael D. White, and Christopher F. Fisher. 2008. Shock value: A comparative analysis of news reports and official police records on TASER deployments. *Policing: An International Journal of Police Strategies and Management*, 31: 148–170.

Roy, O.Z. and Andrew Podgorski. 1989. Tests on a shocking device—The stun gun. *Medical and Biological Engineering and Computing,* 27: 445–448.

Seattle Police Department. 2004. *Seattle Police Department TASER use and deployment fact sheet.* Retrieved November 11, 2008 from taser.com/documents/seattlepduse.pdf.

Smith, Michael R. 2008. Toward a national use-of-force collection system: One small and focused step is better than a giant leap. *Criminology & Public Policy,* 7: 619–627.

Smith, Michael R. and Geoffrey P. Alpert. 2000. Pepper spray: A safe and reasonable response to suspect verbal resistance. *Policing: An International Journal of Police Strategies and Management*, 23: 233–245.

Smith, Michael R., Robert J. Kaminski, Jeffrey Rojek, Geoffrey P. Alpert, and Jason Mathis. 2007. The impact of conducted energy devices and other types of force and resistance on officer and suspect injuries. *Policing: An International Journal of Police Strategies and Management*, 30: 423–446.

Stratbucker, Robert, Richard Roeder, and Max Nerheim. 2003. *Cardiac safety of high voltage Taser X26 waveform*. Proceedings of the 25th Annual International Conference of the IEEE EMBS, Cancun, Mexico.

Strote, Jared, Rich Campbell, John Pease, Shane Hamman, and Range Hutson. 2006. The role of Tasers in police restraint-related deaths. *Annals of Emergency Medicine*, 46: S85.

Surette, Ray. 1998. *Media, crime, and criminal justice: Images and realities*. Belmont, CA: Wadsworth.

Sztajnkrycer, Matthew D. and Amado A. Baez. 2008. *Excited delirium and sudden unexpected death*. Retrieved January 9, 2008 from acep.org.

TASER International. 2006. *TASER non-lethal weapons: Field data as of July 2006*. Retrieved September 23, 2006 from taser.com/facts/documents/Injury_Reduction_ Stats_2Q_2006.ppt.

Tuch, Steven A. and Ronald Weitzer. 1997. The polls-trends: Racial differences in attitudes toward the police. *Public Opinion Quarterly*, 61: 642–663.

Vilke, Gary M. and Theodore C. Chan. 2007. Less lethal technology: Medical issues. *Policing: An International Journal of Police Strategies and Management*, 30: 341–357.

Wake Forest University Baptist Medical Center. 2007. *Press release: Nationwide independent Taser study results suggest devices are safe*. Retrieved October 27, 2007 from 1.wfubmc.ed/ news/articleID=2165.

Walker, Samuel and Charles M. Katz. 2002. *The police in America: An introduction*. New York: McGraw Hill.

White, Michael D. 2002. Identifying situational predictors of police shootings using multivariate analysis. *Policing: An International Journal of Police Strategies and Management,* 25: 726–751.

White, Michael D. 2006. Hitting the target (or not): Comparing characteristics of fatal, injurious, and non-injurious police shootings. *Police Quarterly*, 9: 303–330.

White, Michael D. 2008. Identifying good cops early: Predicting recruit performance in the academy. *Police Quarterly,* 11: 27–49.

White, Michael D. and Justin Ready. 2007. The TASER as a less-lethal force alternative: Findings on use and effectiveness in a large metropolitan police agency. *Police Quarterly*, 10: 170–191.

White, Michael D. and Justin Ready. 2009. The impact of the TASER on suspect resistance: Identifying predictors of effectiveness. *Crime & Delinquency*. Pre-published in OnlineFirst February 26, 2008 DOI 10.1177/0011128707308099.

Wrobeleski, Henry M. and Karen M. Hess. 2003. *Police operations*. Belmont, CA: Wadsworth.

Michael D. White is an associate professor in the School of Criminology and Criminal Justice at Arizona State University. White's primary research interests involve the police, including use of force, misconduct, training, and performance measurement. His recent work has been published in *Criminal Justice and Behavior*, *Journal of Experimental Criminology,* and *Crime & Delinquency.*

Justin Ready is an assistant professor in the School of Criminology and Criminal Justice at Arizona State University. He has conducted research on crime displacement, hot-spot policing strategies, victimology, and surveillance cameras. His articles have appeared in *Justice Quarterly, Crime & Delinquency, Criminology,* and *Police Quarterly.*

POLICY ESSAY

TASERS IN THE MEDIA

Should police departments develop specific training and policies governing use of multiple TASER shocks against individuals who might be in vulnerable physiological states?

Robert J. Bunker
Counter-OPFOR Corporation

This essay focuses on the "policing implications" stemming from the work by White and Ready (2009, this issue). More specifically, it addresses the policy guidance stating:

> *More research is needed to explore the relationship between mental illness, drug use (illicit or therapeutic), continued resistance, and increased risk of death. In the meantime, police departments should develop specific training and policies governing the use of multiple TASER shocks against individuals who may be in these vulnerable physiological and psychological states.*

It is not the intent of this essay to critique or assess the viability of the data set or statistical methodology White and Ready (2009) used to provide this guidance. Enough significant variance in findings and resulting policy implications exists with other studies—most notably, the findings of William Bozeman, lead investigator of a National Institute of Justice (NIJ) study (NIJ, 2008) for others to offer methodological criticisms.[1] For the purposes of this policy essay, we will simply assume that the research findings of the article in question are irrefutable and not subject to debate.

---

1. According to Sid Heal, a subject matter expert on less lethal weapons, newspaper accounts are not a random sample. In fact, there would be something specifically newsworthy about a TASER incident before it would be reported. Hence, such a study makes the most fundamental error in statistics by sampling on the dependent variable. A proposal in to the National Institute of Justice (NIJ) for funding currently seeks to do a definitive study on the subject of the work in question. The data set for that proposed study will draw from $N = {\sim}50,000$ incidents and will include major police departments and two major universities (S. Heal, personal communication, May 5, 2009).

The policy guidance—*police departments should…*—stemming from the work by White and Ready (2009) raises tactical issues for police departments whose personnel use the TASER (TASER International, Inc., Scottsdale, AZ) device. It never was mentioned in the article why police departments should do this, but as stated within the Policy Implications section, it is apparently imperative that they comply. This essay, however, will portray the issues, difficulties, and even pitfalls inherent in using academic (i.e., laboratory and office) research alone to make ensuing policy suggestions—without taking into account actual law-enforcement (i.e., street and field) operations. By employing hypothetical situations, this study examines the tactical issues that could result were this policy guidance actually to be implemented.

The first tactical issue involves law-enforcement officers being required to identify non-compliant individuals encountered who might be in vulnerable physiological and psychological states. Specific states of concern are mental illness and drug use (illicit or therapeutic). This identification requirement would require patrol officers to have an "a priori" knowledge of the physiological and psychological state of noncompliant individuals encountered. In November 2008, the San Antonio Police Department in Texas enacted such a TASER-use policy when it prohibited using the device against suspects "known to be under the influence of drugs." This a priori knowledge is derived from officers observing the noncompliant individual using (drugs) prior to the confrontation per the directive of Chief William McManus who enacted the policy (Chasnoff, 2008). No other a priori guidance has been noted, and expectedly Chief McManus has been raked over the coals in a policing forum (see policeone.com) by rank-and-file officers who want him to get back in touch with the realities of the street. In the chief's defense, for the directive to be in force, officers would be required to view actual drug use by a noncompliant individual. Far more problematic would be a blanket directive such as the one suggested in White and Ready (2009).

Placing the identification burden squarely on the shoulders of officers creates highly ambiguous and, as a result, more stressful situations for them on the street. When faced with a noncompliant and combative individual, TASER-carrying officers will be required—in addition to following normal procedures—to search for signs of drug use and mental illness. This determination will have to be made in the "fog and friction" of the moment with officers having to separate and analyze the signs of public drunkenness, situational mental impairment stemming from rage and other extreme emotions (versus true mental illness), low cognitive abilities, and overt antisocial and defiant behaviors (often exhibited by hardcore gang members) from the vulnerable states of concern. Additionally, per White and Ready (2009), noncompliant individuals shocked in police–citizen encounters are less likely to be healthy. This presumably means that such individuals also will be burdened by one or more adverse physical conditions— such as smoker's cough, hypertension, or asthma—which could become more pronounced because of their struggles and could mask underlying drug use or mental illness conditions. The failure to make an immediate and accurate assessment of the noncompliant individual's vulnerable state in the encounter could leave individual officers open to departmental discipline, censor, and

potentially even civil and criminal liability when toxicological testing is completed or mental health records are secured days after a potentially fatal encounter has passed.

Faced with such unrealistic identification requirements and exposure issues, patrol officers really have only three basic response options at their disposal. The first option is simply not to use the TASER device.[2] Other less lethal technologies such as pepper spray, blunt impact munitions, batons, or old-school approaches such as "pig piles" and possibly even an immediate escalation to lethal options would then be used to fill the void in the policing use-of-force continuum. Because of the high past effectiveness rate and police departments' reliance on the TASER device, however, the injury and death rate of noncompliant individuals, and injuries to officers, surely would rise. The second option is for officers to use one and only one TASER shock against an individual and then to cease their use of the weapon. Were the sole shock not to effectuate an individual's compliance, then officers would fall back to the first option, which includes other less lethal technologies, pig piles, and lethal force. The third option is for officers to approach all encounters with noncompliant individuals as if they exist in one of the vulnerable states of concern. A minority of encounters, composed of 6.3% individuals under the influence of drugs and 22.9% of emotionally disturbed or mentally ill individuals (as per the White and Ready [2009] study), would thus drive officers in their tactical response. For the purposes of this essay, we will continue with the assumption the third option is taken by officers.

This leads us to a second tactical issue, which is derived from navigating the first tactical issue by declaring it insurmountable, and thus, all individuals encountered by officers in a confrontation are now by default assumed to be in a vulnerable physiological or psychological state. This tactical issue focuses on developing the specific training and policies governing the use of multiple TASER shocks against such noncompliant individuals (and, per the assumption above, everyone else encountered too).

Two basic methods can be used to respond to this tactical issue—one is process driven, and the other is issues driven (Rose, 2007). The process-driven method can be viewed in the International Association of Chiefs of Police (IACP) electronic weapons program recommendations (IACP, 2005a). These recommendations are for police departments who seek to select, acquire, and ultimately use TASERS (because of their market-share dominance). Nine steps are highlighted: (1) establishing a leadership team, (2) placing the weapon on the department's use-of-force continuum, (3) assessing costs and benefits, (4) determining staff roles and responsibilities, (5) engaging in community outreach, (6) policy and procedure adoption, (7) training, (8) deploying the weapon in phases, and (9) evaluation and officer compliance. The IACP method would thus promote individual departments to develop their own specific training and policies governing the use of multiple TASER shocks against noncompliant individuals in the vulnerable states. The method does not identify what their policies should be but instead

---

2. This is the case in San Antonio, Texas, with the no-use policy against suspects known to be using drugs still in place. In lieu of TASERS, "Other measures would include other forms of force, whether intermediate or deadly force" (Gabriel Trevino [SAPD PIO]; G. Trevino, personal communication, May 5, 2009).

provides the program template that can be used to create them. As a result, this method provides us with no immediate help in regard to our training and guidelines query.

The second method is issues driven and is promoted by the Police Executive Research Forum (PERF). That organization tends to focus on such items as safety, training, and operational use and has come up with 52 guidelines for conducted energy devices such as TASERS (PERF, 2005). This basically results in a check-off-list approach to policing policy with the intent that police departments will follow these guidelines when using TASERS. These guidelines were not specifically written with individuals who are in vulnerable states of concern in mind, but do tend to err on the side of safety and community perceptions based on the special interests of this group. From a training and guidelines query perspective, some guidelines could be used or modified to help deal with the second tactical issue. Also, multiple IACP Training Keys have been written on TASERS and Electronic Control Weapons (ECWs) with the most relevant one written in 2005 (IACP, 2005b). In addition, a current IACP ECW Model Policy (IACP, 2005c) exists along with an ECW concepts and issues paper (IACP, 2005d).[3] Much of the text is repeated in these documents as they all were created in 2005. Their guidance can be culled for information pertaining to this tactical issue.

The tactical issue at hand is how to approach multiple TASER shocks against drug-using (illicit and therapeutic) and mentally ill individuals. Because these noncompliant individuals did not accede to the threat of a TASER shock or even to an initial TASER shock, it can only be assumed that they will continue to engage in active and continued resistance for an indeterminate amount of time. Once again, patrol officers will have no a priori knowledge of how long a noncompliant individual's struggles will last. Furthermore, the longer their physical exertions last, the greater the chance for excited delirium to come into play because of the individual's vulnerable state. The PERF guidance for operational procedures states the following:

> Use the weapon for one standard cycle and stop to evaluate the situation. If subsequent cycles are necessary, agency policy should restrict the number and duration to the minimum activations necessary to control the situation…. Only one officer should activate a weapon against a person at any time (Rose, 2007: 7).

This guidance does not provide an absolute cap or threshold of TASER shocks or activations that can be used, but rather it calls for the minimum necessary to control the situation (i.e., the de facto obtainment of individual compliance to being handcuffed and placed under arrest). This is somewhat in line with IACP ECW Model Policy (IACP, 2005c: 2), which states the following: "Upon firing the device, the officer shall energize the subject the least number of times and no longer than necessary to accomplish the legitimate operational objective." However, an IACP Training Key guidance from 2005 states the following:

3.   Phil Lynn of the IACP explained that his organization is in the middle of updating their Model Policy on Electronic Weapons based on recent studies and reports (P. Lynn, personal communication, May 5, 2009).

896   *Criminology & Public Policy*

Officers should specifically justify in their use-of-force report any instance in which (1) subjects are energized more than three times, (2) subjects are subjected to any energy cycle longer than 15 seconds in duration, (3) more than one ECW is effectively used against a subject in any given incident, or (4) an ECW is used against a person in a susceptible population group (defined below; IACP, 2005b: 2).

The susceptible population group in this usage is "drug or alcohol use or mental illness" and applies to lessened pain sensitivity vis-à-vis using the TASER by means of a pain compliance technique rather than electromuscular disruption. What is interesting about this guidance is that it creates a three-TASER-shock threshold at which point additional shocks need to be tactically justified for nonsusceptible groups.

A report from British Columbia (BCOPCC, 2005) is also discussed in the Training Key and provides the following guidance:

A single TASER application made before the subject has been exhausted, followed by a restraint technique that does not impair respiration may provide the optimum outcome [*Note—do not use the prone, four-point restraint whenever possible*] (IACP, 2005b: 3).

This guidance refers to situations that involve individuals using stimulant drugs and under alcohol intoxication, and it has been created because of sudden death syndrome potentials. This brings us back to an officer's second basic response option of only one TASER shock use. Here, combined with the appropriate restraint technique, it is now viewed as likely leading to an "ideal" or optimum outcome. Consistently achieving such an ideal on the street, however, is not only highly improbable but also a police administrator's pipe dream.

The implications, then, from the various forms of policing guidance is that to implement a TASER-use policy for those in vulnerable states of concern—and for all other noncomplaint individuals for that matter—we would probably need to create an arbitrary and politically acceptable "firebreak" or "threshold" of TASER shocks allowable in an incident. Given that more research is needed to understand the underlying relationships promoting this need, as discussed in the White and Ready (2009) article, we can do no better than a guesstimate of what this number should be—1, 2, 3, 4, 5…? Failure to create such a TASER shock threshold would mean that police departments would continue to use the TASER in its current manner—that is, compelling a noncompliant individual to stop resisting arrest by means of whatever number of TASER shocks it takes—which would probably be unacceptable. Once the threshold of shocks is reached because a noncompliant individual is still resisting arrest, other less lethal technologies and techniques or lethal force options would then have to be employed to resolve the situation.

Once the noncompliant individual who has been TASER shocked is in custody, then medical attention might be needed. PERF guidelines are:

> Anyone shocked with a weapon be evaluated by a medic and monitored regularly
> while in police custody … officers be made aware of the increased risk of sudden
> death of people under the influence of drugs and "symptoms associated with excited
> delirium" … agencies seek help from medical personnel to develop appropriate
> medical protocols on the weapons and, when possible, notify emergency medical
> personnel when officers respond to calls in which it is anticipated that they may
> use their weapons… (Rose, 2007: 6).

Because drug use and mental illness may not be initially recognized when an individual is taken into custody, it would be best to consider everyone who had been TASER shocked to fall into this high vulnerability group. The IACP Training Key concurs with this assessment:

> Whenever reasonably possible, all individuals who have been incapacitated should
> be transported to an emergency or other medical facility for evaluation. This practice
> takes the guesswork away from the officers, who generally have little information
> on the subject's medical condition or on the effects of the ECW on the subject.
> This protocol provides an element of insurance for the officer and the department
> (IACP, 2005b: 5).

The key goes on to acknowledge that it is by no means practical for all departments, especially small ones in rural areas, to engage in such transport. The key then identifies specific situations that should trigger medical attention being sought, with the following situations being most pertinent:

- When a subject requests medical attention (officers should routinely inquire whether the subject desires medical attention)
- When a subject does not seem to recover properly after being hit (the signs and symptoms of failure to recover from an ECW exposure should be covered in training)
- When the subject is in a potentially susceptible population as defined elsewhere in this document
- When the subject has been energized more than three times
- When a subject has been subjected to a continuous cycle of 15 seconds or more at any time during the incident
- When the subject has exhibited signs of extreme, uncontrollable agitation or hyperactivity prior to ECW exposure, as discussed in the section above on sudden death

Although it can be said that an individual's health and well-being are being considered in this guidance, it can also be viewed as the implementation of lawsuit-avoidance strategies for police departments. The upshot of this guidance is its simplicity: Anyone TASER shocked should probably be transported to some sort of medical facility to be checked out, just in case. No scientific imperative seems to exist for this need—it is just a basic CYA (cover your ass) protocol as a defense against ever-present lawsuits.

What are the lessons learned from this hypothetical attempt at having police departments create specific training and policies governing the use of multiple TASER shocks against individuals who may be in vulnerable physiological and psychological states? They are:

- It will be pretty much impossible for patrol officers to determine which noncompliant individuals are in vulnerable physiological and psychological states and which ones are not.
- Because this tactical issue is insurmountable, officers have three response options: Do not use the TASER, only use one TASER shock, or—if they are going to use the TASER—treat all noncompliant individuals encountered as being in vulnerable physiological and psychological states.
- If option three is chosen, then the creation of an arbitrary and politically acceptable "firebreak" or "threshold" of TASER shocks allowable in an incident will probably be needed. The number three keeps appearing in guidance directives, so this would likely be the arbitrary number chosen.
- Once the threshold of shocks is reached, because a noncompliant individual is still resisting arrest, other less lethal technologies and techniques (holds or pig piles) or lethal force options would then have to be employed to resolve the situation. Individuals on illicit stimulants (e.g., crack, PCP, and methamphetamine) and who have severe mental illness can be extremely dangerous, so officer safety concerns will have to now be a priority.
- Because the newer TASERs typically have an extremely high effectiveness rating from law enforcement (nearly unheard of for any force option, including lethal force; Bovbjerg, Heal, and Kenny, 2009), the fallback to less effective, less lethal technologies, techniques, and lethal force options for the still unresolved incidents would presumably result in more noncompliant individual injuries and deaths and may even result in additional officer injuries. This is ultimately contrary to the public good because more (rather than less) people will be injured and killed, but it is a logical outcome from implementation of the above.
- After all incidents in which a noncompliant individual has been TASER shocked, to cover anyone even remotely liable to be in a vulnerable physiological and psychological state, the shocked individual should now be transported to a medical emergency room or facility for evaluation. No scientific basis exists for the need for such widespread transport, but it is ultimately a prudent lawsuit-avoidance strategy for a police department.[4]

As can be observed from the hypothetical implementation of the policy guidance from a street or tactical level, it would be viewed as pretty much a disaster. The transition from academic research to policy suggestions or guidance to actual law-enforcement operations in the field ultimately breaks down in this instance. The street cop perspective, "if you take away my TASER, or severely curtail how I may use it, then what is the alternative tool you will give me to do my job?" thereby comes in to play. Sid Heal, retired from the Los Angeles Sheriff's

4. Some inquiry into the costs of unnecessary transport and evaluation should be made, especially when weighted against opportunity costs as a percentage of police budgets.

Department, and two other writers have estimated that even in a worst-case scenario in which 291 people died from TASER-related deaths (Amnesty International. 2007: 2) out of the 70,000 TASER-shock incidents (General Accounting Office, 2005: 1) with noncompliant individuals that have taken place, the death rate would be at most in the region of 1 in 4,000 incidents. This is compared with the chance of death during the use of lethal force in similar situations, which is 1 in 4 incidents (Bovbjerg et al., 2009). Until you can provide another less lethal device with a better field-use record (or, in the case of a new innovation, promising field-use potentials) than the TASER, suggestions to replace it with other weapons, tools, and techniques—including falling back to the use of lethal force—should be tempered with the realities of the street and the basic tactical issues that such suggestions may trigger.

References
Amnesty International. 2007. *Amnesty International's concerns about TASER® use: Statement to the U.S. Justice Department inquiry into deaths in custody*. London, UK: Author.

Bovbjerg, Viktor, Sid Heal, and John Kenny. 2009. Voodoo science and non-lethal weapons. *Tactical Response*, 24–26.

British Columbia Office of the Police Complaint Commissioner (BCOPCC). 2005. *TASER technology review final report*. Victoria, British Columbia: Author.

Chasnoff, Brian. 2008. Texas chief prohibits TASER use against suspects on drugs. *San Antonio Express-News*, October 15. Retrieved May 4, 2009 from PoliceOne.com.

Government Accountability Office. 2005. *TASER weapons: Use of Tasers by selected law-enforcement agencies*. Washington, DC: Author.

International Association of Chiefs of Police (IACP). 2005a. *Electro-muscular disruption technology (EMDT): A nine-step strategy for effective deployment*. Executive Brief. Alexandria, VA: Author.

International Association of Chiefs of Police (IACP). 2005b. *Electronic control weapons (ECWs): Update 2005*. Training Key 583. Alexandria, VA: Author.

International Association of Chiefs of Police (IACP). 2005c. *Electronic control weapons model policy*. Alexandria, VA: Author.

International Association of Chiefs of Police (IACP). 2005d. *Electronic control weapons*. Alexandria, VA: Author.

National Institute of Justice (NIJ). 2008. *Study of deaths following electro muscular disruption: Interim report*. Washington, DC: Office of Justice Programs, U.S. Department of Justice.

Police Executive Research Forum (PERF). 2005. *PERF conducted energy device policy and training guidelines for consideration*. Washington, DC: PERF Center on Force and Accountability.

Rose, Veronica. 2007. *Taser use guidelines*. OLR Research Report 2007-R-0068. Hartford: State of Connecticut General Assembly.

White, Michael D. and Justin Ready. 2009. Examining fatal and nonfatal incidents involving the TASER: Identifying predictors of suspect death reported in the media. *Criminology & Public Policy*. This issue.

Robert J. Bunker is the CEO of the Counter-OPFOR Corporation. His less lethal weapons (LLW) and policing operations experience includes being a past consultant to the National Law Enforcement and Corrections Technology Center—West (NLECTC—West), a National Institute of Justice (NIJ) grant funded program, a past member of the NIJ Less Lethal Weapons Working Group, numerous publications on LLW, the author of International Association of Chiefs of Police (IACP) Training Keys, the author of tactical-level law-enforcement focused articles, counterterrorism operations and red teaming activities, current participation with a Police Futurists International (PFI) and Federal Bureau of Investigation (FBI) working group on policing futures, and membership in the National Tactical Officers Association (NTOA). He holds six university degrees, including a Ph.D. in political science with an area of concentration in public policy.

Research on conducted energy devices

Findings, methods, and a possible alternative

Robert J. Kaminski
University of South Carolina

The widespread adoption of oleoresin capsicum (OC) or "pepper spray" in the 1990s and the recent widespread adoption of advanced conducted energy devices (CEDs), such as the TASER (TASER International, Inc., Scottsdale, AZ) and STINGER (Stinger Systems, Inc., Tampa, FL), by law-enforcement agencies has generated substantial controversy and public debate. Many of the concerns regarding CEDs parallel those concerns regarding OC (ACLU, 1993, 1995; Amnesty International, 2008; Kaminski, 2005). Although controversy surrounding the use of OC by police largely has subsided, controversy surrounding CEDs continues unabated. One major ongoing concern is whether CED exposure contributes to or causes death. In response, research activity on CEDs has increased in recent years (e.g., Kroll and Ho, 2009). Although design limitations make definitive conclusions regarding the causal or contributory role of CEDs in unexpected in-custody deaths difficult, the preponderance of the evidence suggests that the risk of death or serious injury from CED exposure is low (Council on Science and Public Health, 2009). Still, their safety profile continues to be a matter of debate.

In their contribution to this literature, Michael D. White and Justin Ready (2009, this issue) compare media accounts of fatal and nonfatal CED incidents to identify potential correlates of CED-proximate deaths. Among their stated policy implications, White and Ready call for additional research on the correlates they identified as being associated with the risk of death due to CED exposure (e.g., mental illness, drug use, and continued physical exertion). This policy essay responds to their call for additional research on the correlates and causes of CED-proximate deaths. Specifically, I provide an updated literature review, highlight some limitations of the extant research, and suggest an alternative methodology for developing estimates of the relative risk of death associated with CED exposure.

Direct correspondence to Robert J. Kaminski, Department of Criminology and Criminal Justice, University of South Carolina, Columbia, SC 29208 (e-mail: kaminskb@gwm.sc.edu).

Research Approaches to the Study of CED-Proximate Deaths and Injuries

Several methodological approaches have been employed to study the risk of death and injury associated with CED exposure. For the purposes of this essay, these approaches are classified as controlled trials, quasi-experiments, correlational studies, and descriptive studies.

Controlled Trials

A major concern regarding police deployment of CEDs is whether exposure can cause ventricular fibrillation (VF) in humans. To answer this question, medical researchers have conducted several controlled trials in which animals (pigs, dogs, or sheep) or healthy human subjects were exposed to CED discharges of various intensities and duration. Animal-based studies have found that greater output discharges (i.e., 15–20 times the standard) or discharges of longer duration (e.g., two 40-second exposures) could induce VF or increased heart rhythm and—in rare instances—death in some pigs, but standard discharges of relatively short duration (i.e., 5–15 seconds) produced no VF (Dennis et al., 2007; Esquivel, Dawe, Sala-Mercado, Hammond, and Bir, 2007; Ho, Miner, Lakireddy, Bultman, and Heegaard, 2006; Lakireddy et al., 2008; McDaniel, Stratbucker, Nerheim, and Brewer, 2005; Nanthakumar et al., 2006; Roy and Podgorski, 1989; Stratbucker, Roeder, and Nerheim, 2003; Walter et al., 2008). In some studies, CED barbs were oriented across the hearts of pigs (to simulate a worst-case scenario of a current vector that directly passes through the heart) using standard discharges of up to 15 seconds. Although these studies have found stimulation of the heart muscle, three observed no VF (Lakireddy et al., 2006; Nanthakumar et al., 2006; Roy and Podgorski,1989), and a fourth observed two instances of VF (but no deaths) for four pigs exposed to twenty-seven 10-second shocks (Valentino et al., 2008).

Controlled trials using volunteer human subjects have observed increases in heart rate but no VF after exposure to standard discharges of up to 20 seconds (Bozeman, Barnes, Winslow, Johnson, Phillips, and Alson, 2009; Ho et al., 2008a; Levine, Sloane, Chan, Dunford, and Vilke, 2007). In other studies, no VF was induced in subjects exposed to two or three simultaneous 5-second exposures (Ho, Dawes, and Miner, 2009) or in subjects exposed to 10-second discharges directly over the heart (Ho, Dawes, Reardon, Lapine, and Miner, 2008b). Additional studies have evaluated the effects of CED exposures of up to 20 seconds on respiration (Dawes, Ho, Johnson, Lundin, and Miner, 2007a, 2007b; Dawes, Ho, and Minor, 2008; Ho, Johnson, and Dawes, 2007), blood chemistry (Dawes et al., 2007a, 2007b; Ho et al., 2006; Sloane et al., 2008; Vilke et al., 2007), and core body temperature (Dawes et al., 2007a, 2008). These studies have generally found no adverse effects.

A major benefit of controlled trials is that they are stronger in internal validity and thus useful for establishing causality, but the trials' "Achilles' heel" is that they are weak on external validity, and relationships observed in artificial settings might not hold in real-world conditions. To enhance external validity, several controlled trials have more closely emulated field conditions within the laboratory. Lakireddy et al. (2006) shocked five pigs before and after infusions of

cocaine. Surprisingly, the results indicated that cocaine was protective for CED-induced VF. Ho, Dawes, Cole, et al. (2009) administered varying doses of methamphetamine to 16 sheep that received four exposures each of durations up to 45 seconds. No VF was observed in any of the animals. Nanthakumar et al. (2006) exposed four pigs infused with adrenalin (to simulate physiologic stress) to 16 shocks across the heart; one exposure produced VF.

In a study of human volunteers, eight subjects received a 5-second shock after rigorous exercise (to simulate a physical struggle or flight). No clinically significant or lasting changes in cardiovascular levels were found (Vilke et al., 2007). Other researchers have induced various physiologic states in human subjects, including acidosis, exercise-induced exhaustion, and alcohol intoxication. No significant negative impacts on blood acidosis levels, respiration, or cardiac function were observed after exposure (Ho et al., 2007; Ho, Dawes, Bultman et al., 2009; Moscati et al., 2007).

These studies suggest that CEDs are relatively safe when used on healthy at-rest individuals and on those individuals subject to certain physiological stresses under controlled conditions. However, because controlled trials cannot simulate field conditions completely, they ultimately cannot tell us with certainty whether CEDs cause or contribute to unexpected in-custody death.

Correlational Studies

Correlational designs are those studies in which an assumed cause-and-effect relationship is specified but other experiment features (e.g., pre-tests and control groups) are lacking. They are often cross-sectional, rely on statistical controls for potential rival explanations, and typically cannot support strong causal inferences (Shadish, Cook, and Campbell, 2002). A few correlational studies have used regression analyses to examine the relationship between CED use and officer and suspect injuries; these studies have generally found that CED use was associated with fewer injuries to officers and suspects and with less severe injuries to suspects (MacDonald, Kaminski and Smith, 2009; Smith, Kaminski, Alpert, Fridell, MacDonald, and Kubu, 2009; Smith, Kaminski, Rojek, Alpert, and Mathis, 2007). However, because of their rarity, these studies did not examine deaths.

In another correlational study, White and Ready (2009) conducted a national search of media accounts of fatal and nonfatal CED incidents. Using logistic regression in addition to other methods, they identified several situational and personal characteristics that were associated with accounts of CED-proximate deaths. Among other findings, the number of CED discharges was unrelated to the odds of death, alcohol intoxication was inversely related, and perceived drug impairment and mental illness were positively related to the odds of death. This study is a good initial step in the search for factors that might explain why some exposed suspects die whereas others survive, but—as with other correlational studies—it is a weak design for causal inference and the results should be interpreted with caution. An additional limitation is that

only a small fraction of nonfatal CED incidents were reported in the media during the 5-year study period, and the events analyzed might not be typical.[1]

Descriptive Studies

In epidemiology, descriptive studies typically are conducted early in the phase of a disease investigation when little is known about its frequency or causes. Such investigations explore and describe the general disease patterns in the population, and findings are used to generate hypotheses about the disease's causes and prevention, among other things (Gertsman, 1998). Several descriptive studies of CEDs have been conducted.

Eastman et al. (2007) conducted a prospective study of 426 CED exposures from November 2004 to January 2006 involving the Dallas Police Department. All exposed suspects received medical evaluation. Nonfatal injuries were all minor—small bruises, abrasions, and lacerations that did not require special treatment. One fatality occurred, although White and Ready (2009) concluded that the subject likely would have died even without CED exposure.[2]

In a second prospective study, Bozeman et al. (2009b) examined police and medical records of 1,201 suspects shocked by CEDs in six agencies from June 2005 through June 2008. All exposed suspects received pre-incarceration medical screening. Overall, 1,198 (99.75%) suspects experienced no injuries or mild injuries (i.e., primarily superficial puncture wounds from CED darts), but 3 suspects (0.25%) suffered significant injuries (i.e., head injuries sustained in falls and rhabdomyolysis, a rapid breakdown of skeletal muscle tissue). Two additional suspects died unexpectedly in police custody. According to Bozeman et al. (2009), the relationship between rhabdomyolysis and CED exposure was unclear and the two deaths were determined to be unrelated to CED exposure following autopsy.

Three retrospective case series studies have been conducted of autopsy and toxicology reports of deaths proximate to CED exposure (Kornblum and Reddy, 1991; Strote and Hutson, 2006; Swerdlow, Fishbein, Chaman, Lakkireddy, and Tchou, 2009). The researchers found that many subjects were under the influence of drugs, were behaving in a bizarre manner, or had a pre-existing cardiovascular disease at the time of their death. The general conclusion was that CEDs were not a common cause or contributor to sudden in-custody death.

Findings from the prospective descriptive studies suggest that most CED-related injuries are minor and that CED-proximate deaths seem to be rare relative to the number of exposures (i.e., less than 0.25% in both studies). The retrospective studies of CED-proximate deaths indicate that most CED-exposed subjects were under the influence of drugs, in a highly agi-

---

1.   For example, data from 12 law-enforcement agencies of various sizes showed 5,470 CED uses (Smith et al., 2009). Because many thousands of law-enforcement agencies adopted CEDs between 2002 and 2006, the 333 uses reported by the media necessarily represent a tiny fraction of the total amount. Furthermore, an additional 30 fatal CED incidents during the study period were not accounted for (see Amnesty International, 2008).

2.   The suspect became unresponsive sometime after being shocked and restrained. EMS transported the subject to a hospital with a core body temperature of 107.4 °F. He died shortly after arrival. The medical examiner later reported high levels of cocaine in the subject's bloodstream.

tated state (possibly excited delirium), suffered from heart disease, or were under restraint by law-enforcement personnel. It seems, therefore, that many fatal encounters in which CEDs are used involve subjects already at risk for sudden death from other causes.

A strength of descriptive epidemiological studies is their utility in providing assessments of the safety of CEDs as used by law-enforcement officers in real-world conditions. Their use is limited, however, for making causal assessments. Furthermore, the prospective studies highlight the difficulties and cost of obtaining data on statistically rare events. In Bozeman et al.'s (2009) study, for example, 3 moderate-to-severe injuries and 2 deaths occurred out of 1,201 exposures.

Quasi-Experiments

Quasi-experiments are distinguished from true experiments primarily by lack of random assignment of subjects to experimental and control groups, but they still are useful for testing causal hypotheses when random assignment is not feasible (Shadish et al., 2002). Few quasi-experiments have been conducted on CEDs outside of laboratory settings.

Using 108 months of pre–post data from the Orlando Police Department and 60 months of pre–post data from the Austin Police Department in Texas, MacDonald et al. (2009; see also Smith et al., 2009) tested the effect of CED adoption on monthly rates of officer and suspect injuries in two time-series regression analyses and found substantial reductions in both jurisdictions. Although interrupted time-series models rule out many threats to internal validity, they can be strengthened by incorporating one or more control sites, including one or more nonequivalent dependent variables, or by adding other design features (Britt, Kleck, and Bordua, 1996; Shadish et al., 2002). Unfortunately, the MacDonald et al. study could not examine deaths.

In a study designed to test the effect of CED deployment on rates of sudden, in-custody death in the absence of lethal force, Lee et al. (2009) obtained data for each of the 5 years preceding CED deployment through the 5 years succeeding CED deployment from 50 (40%) of the 126 agencies surveyed. Controlling for the arrest rate, Lee et al. found that the rate of in-custody sudden death increased 6.4 times in the first full year after deployment compared with the average rate in the 5 years before deployment. The authors speculated that high initial rates of CED use contributed to the increase in sudden deaths by escalating some confrontations to the point that officers needed to resort to deadly force.[3] Limitations of this study (see O'Riordan, 2009) include the low response rate and the fact that only 40 of the 50 responding agencies provided data on deaths in the first year after CED deployment, which could have biased the findings. Furthermore, the survey did not determine whether subjects who died were actually exposed to a CED. An additional concern is that there were likely very few deaths

3. Lee et al.'s (2009) speculation is implausible on its face. Rather, given the deterrent and incapacitative effects of CEDs, it is much more likely that they end many confrontations before they escalate into deadly force situations (Kaminski, Edwards, and Johnson, 1998; Scharf and Binder, 1993).

(counts were not reported), and so the increase in deaths in the year after CED deployment could have occurred by chance.

Case-Control Studies—A Potentially Useful Alternative

Although additional designs beyond those described above for assessing the risk of death or serious injury from CED exposure are possible (e.g., prospective or retrospective cohort studies), one particularly attractive method is the case-control study. Case-control studies have been used widely in epidemiology to estimate the effect of an exposure on the risk of a disease (or other outcome; Gerstman, 1998; Schlesselman, 1982). They also have been usefully employed in criminology (e.g., Dobrin, 2001; Loftin and McDowall, 1988). Case-control studies are especially appropriate when the outcome of interest is rare. Although cohort studies provide several advantages compared with case-control studies, they are inefficient for analyzing rare events and they are relatively more expensive and time consuming: Typically, they require large samples and long follow-up periods to obtain sufficient numbers of cases for analysis (Gerstman, 1998; Schlesselman, 1982).

Case-control studies are used to identify a factor or factors that contribute to, or potentially cause, some outcome of interest by comparing subjects who have experienced the outcome (the case group) with subjects who did not experience the outcome but are otherwise similar (the control group). The frequency of exposure to the risk factor of interest (e.g., CED shock) among the cases is then compared with the frequency of exposure among the controls, which allows for an assessment of whether exposure to the risk factor significantly contributed to the outcome (e.g., in-custody death; Goodman, Mercy, Layde, and Thacker, 1988).

An initial step in a case-control study is the identification and selection of subjects who experienced the event of interest.[4] This group might consist of all unintentional deaths (or a random sample thereof) from a predefined population. If the outcome of interest is rare, the search for cases might need to be extended to several jurisdictions and further backwards in time, which can pose significant challenges (Schlesselman, 1982). This issue is likely to be the case for unexpected deaths in police custody, because they generally seem to be rare occurrences (see Mumola, 2007).

Controls must not have experienced the event of interest (i.e., death), and they must be as similar as possible to the cases in terms of the potential for past exposure to the suspected cause and other factors (Schlesselman, 1982). Thus, for each case identified in any particular law-enforcement agency, one or more controls might be selected at random from the same agency (or matched on various characteristics). Because controls should be similar to cases in terms of "eligibility" for CED exposure, they could be selected from among suspects who actively resisted police officers.

---

4.   One possible approach is to use the list of CED-proximate deaths compiled by Amnesty International (2008b), which identifies the involved agencies. The agencies or a subset could be contacted and asked to provide data on the cases and a sample of controls. (For an additional source of deaths, see truthnottasers.blogspot.com/2008/04/what-follows-are-names-where-known.html.)

This approach has several benefits. First, the application of the multivariate logistic regression model to case-control data is straightforward, and it allows for an adjustment of confounding variables that could account for or distort an association between an exposure and an outcome variable (Goodman et al, 1988; Loftin and McDowall, 1988). Second, interaction terms between potential risk factors can be included (Schlesselman, 1982). These factors are important considerations, because it is frequently unclear whether suspects die unexpectedly in custody from preexisting illness, drug intoxication, exertion, physical restraint, CED exposure, or some combination of these or other factors. Third, although logistic regression analysis of case-control data can produce inflated effect sizes (Liberman, 2005), when the outcome of interest is rare, case-control studies provide good estimates of the relative risk in the target population (Loftin and McDowall, 1988; Schlesselman, 1982).

No method employed to date allowed for an analysis of CED exposure and risk of death while controlling for other factors or conditions commonly observed in the field. Controlled trials are limited, because they cannot completely simulate field conditions. White and Ready (2009), consistent with the descriptive epidemiological studies reviewed earlier, did not include an unexposed comparison group, thus precluding estimates of relative risk. The correlational and quasi-experimental studies by MacDonald et al. (2009) and Smith et al. (2007, 2009)—although useful for examining the effects of CEDs on (mostly minor) injuries—could not analyze deaths. The quasi-experiment by Lee et al. (2009) included the number of arrests as an offset in their regression model, but controlled for no other potential confounders. Furthermore, the number of people who were exposed *and* experienced the outcome is unknown, and it is possible that none or few of the suspects who died were exposed to a CED. Causal inference is, therefore, weak.

The case-control design has its own limitations, and its successful application depends on several factors (see Schlesselman, 1982). One of the major challenges will be obtaining reliable and valid data on cases and controls (Smith, 2008). However, this difficulty should be offset to some degree because case-control designs typically involve data collection on relatively few subjects (Schlesselman, 1982). If it is carried out successfully, then a case-control study would provide valuable information in regard to the contributors to in-custody death.

References

ACLU of Southern California. 1993. *Pepper spray: A magic bullet under scrutiny.* Los Angeles: Author.

ACLU of Southern California. 1995. *Pepper spray update: More fatalities, more questions.* Los Angeles: Author.

Amnesty International. 2008. *List of deaths following use of stun weapons in U.S. law enforcement.* London, UK: Author.

Bozeman, William P., William E. Hauda, Joseph J. Heck, Derrel D. Graham, Brian P. Martin, and James E. Winslow. 2009. Safety and injury profile of conducted electrical weapons used by law enforcement officers against criminal suspects. *Annals of Emergency Medicine,* 53: 480–489.

Britt, Chester L., Gary Kleck, and David J. Bordua. 1996. A reassessment of the D.C. gun law: Some cautionary notes on the use of interrupted time series designs for policy impact assessment. *Law & Society Review,* 30: 361–380.

Council on Science and Public Health. 2009. *Use of Tasers® by law enforcement agencies.* Report 6 of the Council on Science and Public Health. Chicago, IL: American Medical Association. Retrieved August 3, 2009 from taser.com/research/science/Pages/default.aspx.

Dawes, Donald M., Jeffery D. Ho, Mark A. Johnson, Erik J. Lundin, and James R. Miner. 2007a. 15-second conducted electrical weapon application does not impair basic respiratory parameters, venous blood gases, or blood chemistries and does not increase core body temperature. *Annals of Emergency Medicine,* 50: S6.

Dawes, Donald M., Jeffery D. Ho, Mark A. Johnson, Erik J. Lundin, and James R. Miner. 2007b. Breathing parameters, venous blood gases, and serum chemistries with exposure to a new wireless projectile conducted electrical weapon in human volunteers. *Annals of Emergency Medicine,* 50: S133.

Dawes, Donald M., Jeffery D. Ho, and James R. Miner. 2008. The effect of a cross-chest electronic control device exposure on breathing. *Annals of Emergency Medicine,* 54: S65.

Dennis, Andrew J., Daniel J. Valentino, Robert J. Walter, Kimberly K. Nagy, Jerry B. Winners, Faran Bokhari, et al. 2007. Acute effects of TASER X26 discharges in a swine model. *Journal of Trauma, Injury, Infection, and Critical Care*, 63: 581–590.

Dobrin, Adam. 2001. The risk of offending on homicide victimization: A case control study. *Journal of Research in Crime and Delinquency,* 38: 154–173.

Eastman, Alexander L., Jeffrey C. Metzger, Fernando L. Benitez, James Decker, Kathy J. Rinnert, Paul E. Pepe, et al. 2007. Conductive electrical weapons: A prospective, population-based study of the safety of application by law enforcement. *Journal of Trauma, Injury, Infection, and Critical Care,* 62: 265–275.

Esquivel, Amanda O., Elisabeth J. Dawe, Javier A. Sala-Mercado, Robert L. Hammond, and Cynthia A. Bir. 2007. The physiological effects of a conducted electrical weapon in swine. *Annals of Emergency Medicine,* 50: 576–583.

Gerstman, Burt B. 1998. *Epidemiology kept simple: An introduction to classic and modern epidemiology.* New York: Wiley.

Goodman, Richard A., James A. Mercy, Peter M. Layde, and Stephen B. Thacker. 1988. Case-control studies: Design issues for criminological applications. *Journal of Quantitative Criminology,* 4: 71–84.

Kroll, Mark W. and Jeffery D. Ho. 2009. *TASER® conducted electrical weapons: Physiology, pathology, and law.* New York: Springer.

Ho, Jeffrey D., Donald M. Dawes, Laura L. Bultman, Ronald M. Moscati, Timothy A. Janchar, and James R. Miner. 2009. Prolonged TASER use on exhausted humans does not worsen markers of acidosis. *American Journal of Emergency Medicine,* 27: 413–418.

Ho, Jeffrey D., Donald M. Dawes, Jon B. Cole, Robert F. Reardon, Julie C. Hottinger, Karen S. Terwey, et al. 2009. *Effect of an electronic control device exposure on a methamphetamine intoxicated animal model.* Paper presented at the Society for the Academy of Emergency Medicine Scientific (SAEM) Assembly, New Orleans, LA.

Ho, Jeffrey D., Donald M. Dawes, and James R. Miner. 2009. *Multiple simultaneous exposures of electronic control devices (ECDs) in human volunteers.* Hennepin County Medical Center, Minneapolis, MN. Retrieved July 31, 2009 from taser.com/research/Science/Pages/CurrentElectronicControlDeviceResearch.aspx.

Ho, Jeffrey D., Donald M. Dawes, Robert F. Reardon, Anne L. Lapine, Benjamin J. Dolan, Erik J. Lundin, et al. 2008a. Echocardiographic evaluation of a TASER-X26 application in the ideal human cardiac axis. *Academic Emergency Medicine,* 15: 838–844.

Ho, Jeffrey D., Donald M. Dawes, Robert F. Reardon, Anne L. Lapine, and James R. Miner. 2008b. Echocardiographic determination of cardiac rhythm during trans-thoracic wireless conducted electrical weapon exposure. *Annals of Emergency Medicine,* 52: S62.

Ho, Jeffrey D., Donald M. Dawes, Frank J. Ryan, Erik J. Lundin, Kenneth G. Overton, Aadam J. Zeiders, et al. 2009. *Catecholamines in simulated arrest scenarios.* Paper presented at the Australasian College of Emergency Medicine Winter Symposium, West Melbourne, Victoria, Australia.

Ho, Jeffrey D., Mark A. Johnson, and Donald M. Dawes. 2007. *The state of current human research and electronic control devices (ECDs).* Paper presented at the 4th European Symposium on Non-Lethal Weapons, Ettlingen, Baden-Wurttemberg, Germany.

Ho, Jeffrey D., James R. Miner, Dhanunjaya R. Lakireddy, Laura L. Bultman, and William G. Heegaard. 2006. Cardiovascular and physiologic effects of conducted electrical weapon discharge in resting adults. *Academic Emergency Medicine,* 13: 589–595.

Kaminski, Robert J. 2005. *Common issues regarding oleoresin capsicum and Tasers.* Presented at the U.S. Department of Justice On-line Symposium on Less Lethal Force.

Kaminski, Robert J., Steven M. Edwards, and James W. Johnson. 1998. The deterrent effects of oleoresin capsicum on assaults against police: Testing the velcro-effect hypothesis. *Police Quarterly*, 1: 1–20.

Kornblum, Ronald N. and Srinevas K. Reddy. 1991. Effects of the TASER in fatalities involving police confrontation. *Journal of Forensic Sciences*, 36: 434–438.

Kroll, Mark W. and Jeffrey D. Ho. 2009. *TASER® conducted electrical weapons: Physiology, pathology, and law.* New York: Springer.

Lakkireddy, Dhanunjaya, Donald Wallick, Kay Ryschon, Mina K. Chung, Jagdish Butany, David Martin, et al. 2006. Effects of cocaine intoxication on the threshold of stun gun induction of ventricular fibrillation. *Journal of the American College of Cardiology,* 48: 805–811.

Lakkireddy, Dhanunjaya, Donald Wallick, Atul Verma, Kay Ryschon, William Kowalewski, Oussama Wazni, et al. 2008. Cardiac effects of electrical stun guns: Does position of barbs contact make a difference? *Pacing and Clinical Electrophysiology,* 31: 398–408.

Lee, Bryon K., Eric Vittinghoff, Dean Whiteman, Minna Park, Linda L. Lau, and Zian H. Tseng. 2009. Relation of Taser (electrical stun gun) deployment to increases in in-custody sudden deaths. *American Journal of Cardiology,* 103: 877–880.

Levine, Saul D., Christian M. Sloane, Theodore C. Chan, James V. Dunford, and Gary M. Vilke. 2007. Cardiac monitoring of human subjects exposed to the TASER®. *Journal of Emergency Medicine,* 33: 113–117.

Liberman, Akiva M. 2005. How much more likely? The implications of odds ratios for probabilities. *American Journal of Evaluation,* 26: 253–266.

Loftin, Colin and David McDowall. 1988. The analysis of case-control studies in criminology. *Journal of Quantitative Criminology,* 4: 85–98.

MacDonald, John M., Robert J. Kaminski, and Michael R. Smith. 2009. The effect of less-lethal weapons on injuries in police use of force events. *American Journal of Public Health.* In press.

McDaniel, Wayne C., Robert A. Stratbucker, Max Nerheim, and James E. Brewer. 2005. Cardiac safety of neuromuscular incapacitating defensive devices. *Pacing and Clinical Electrophysiology,* 28: 284–287.

Moscati, Ronald M., Jeffrey D. Ho, Donald M. Dawes, James Miner, Robert Reardon, William Heegaard, et al. 2007. Physiologic effects of prolonged conducted electrical weapon discharge on intoxicated adults. *Academic Emergency Medicine,* 15: S63.

Mumola, Christopher. 2007. *Arrest-related deaths in the United States, 2003–2005.* Washington, DC: Bureau of Justice Statistics.

Nanthhakumar, Kumaraswamy, Ian M. Billingsley, Stephane Masae, Paul Dorian, Douglas Cameron, Vijay S. Chauhan, Eugene Downar, and Elias Sevaptsidis. 2006. Cardiac electrophysiological consequences of neuromuscular incapacitating device discharges. *Journal of the American College of Cardiology,* 48: 794–804.

O'Riordan, Michael. 2009. Adoption of stun guns spikes the risk of in-custody death in the first year. *Medscape Today.* Retrieved July 30, 2009 from medscape.com/viewarticle/587624.

Roy, Orest Z. and Andrew S. Podgorski. 1989. Tests on a shocking device—The stun gun. *Medical and Biological Engineering and Computing,* 27: 445–448.

Scharf, Peter and Arnold Binder. 1993. *The badge and the bullet: Police use of deadly force.* Westport, CT: Praeger.

Schlesselman, James J. 1982. *Case-control studies: Design, conduct, analysis.* New York: Oxford University Press.

Shadish, William R., Thomas D. Cook, and Donald T. Campbell. 2002. *Experimental and quasi-experimental designs for generalized causal inference.* Boston, MA: Houghton Mifflin.

Sloane, Christian M., Theodore C. Chan, Saul D. Levine, James V. Dunford, Tom S. Neuman, and Gary M. Vilke. 2008. Serum troponin I measurement of subjects exposed to the Taser X-26. *Journal of Emergency Medicine,* 35: 29–32.

Smith, Michael. 2008. Toward a national use-of-force data collection system: One small (and focused) step is better than a giant leap. *Criminology & Public Policy,* 7: 619–627.

Smith, Michael R., Robert J. Kaminski, Geoffrey P. Alpert, Lorie A. Fridell, John MacDonald, and Bruce Kubu. 2009. *A multi-method evaluation of police use of force outcomes.* Washington, DC: National Institute of Justice.

Smith, Michael R., Robert J. Kaminski, Jeffery Rojek, Geoffrey P. Alpert, and Jason Mathis. 2007. The impact of conducted energy devices and other types of force and resistance on police and suspect injuries. *Policing: An International Journal of Police Strategies and Management*, 30: 443–426.

Stratbucker, Robert A., Rebecca Roeder, and Max Nerheim. 2003. Cardiac safety of high voltage TASER X26 waveform. *Proceedings from the Engineering in Medicine and Biology Society Annual International Conference,* 4: 3261–3262.

Strote, Jared and H. Range Hutson. 2006. TASER use in restraint-related deaths. *Prehospital Emergency Care,* 10: 447.

Swerdlow, Charles D., Michael C. Fishbein, Linda Chaman, Dhanunjaya R. Lakkireddy, and Patrick Tchou. 2009. Presenting rhythm in sudden deaths temporally proximate to discharge of TASER conducted electrical weapons. *Academic Emergency Medicine,* 16: 726–739.

Valentino, Daniel J., Robert J. Walter, Andrew J. Dennis, Bosko Margeta, Frederic Starr, Kimberly K. Nagy, et al. 2008. Taser X26 discharges in swine: Ventricular rhythm capture is dependent on discharge vector. *Journal of Trauma Injury, Infection, and Critical Care,* 65: 1478–1485.

Vilke, Gary M., Christian M. Sloane, Katie D. Bouton, Fred W. Kolkhorst, Saul D. Levine, Tom S. Neuman, et al. 2007. Physiological effects of a conducted electrical weapon on human subjects. *Annals of Emergency Medicine,* 26: 1–4.

Walter, Robert J., Andrew J. Dennis, Daniel J. Valentino, Bosko Margeta, Kimberly K. Nagy, Faran Bokhari, et al. 2008. TASER X26 discharges in swine produce potentially fatal ventricular arrhythmias. *Academic Emergency Medicine,* 15: 66–73.

White, Michael D. and Justin Ready. 2009. Examining fatal and nonfatal incidents involving the TASER: Identifying predictors of suspect death reported in the media. *Criminology & Public Policy.* This issue.

Robert J. Kaminski is an associate professor in the Department of Criminology and Criminal Justice at the University of South Carolina. He received his doctorate in criminal justice from the University at Albany. Prior to joining the faculty at the University of South Carolina, he held analyst positions with the National Institute of Justice and the New York City Criminal Justice Agency. His research focuses on policing issues, primarily violence against the police, police use of force, and less lethal technologies.

Conducted energy weapons

Learning from operational discretion and encounter outcomes

Gregory B. Morrison

Ball State University

Many interesting aspects are presented in Michael D. White and Justin Ready's (2009, this issue) study that examines media reports of police use of conducted energy devices (CEDs). In my essay, I will focus on three topics: first, the implications of media interest for studying CED deployments; second, police CED policy and training; and third, the much-needed collaboration between governmental and nongovernmental organizations that will be central for future advances in officer performance and public safety when police reasonably must use force.

TASER International, Inc. (Scottsdale, AZ) introduced its X26 in 2003, and this make and model is what most springs to mind upon hearing or reading the word "TASER"[1] (TASER International, Inc., 2009). Compared with the company's previous M26, the X26 is sleek, small, as well as light and can be carried in a plastic holster on the duty belt. The handiness of the X26, when combined with its demonstrated effectiveness to incapacitate suspects,[2] has made it the standard tool against which other CEDs are compared.

Dramatically increased media interest in police use of CED technology is evident because half (48.9%) of the media reports that White and Ready (2009) gathered for their 2002–2006 study period came from 2006 alone. Many departments have been using CEDs for several years and relatively regularly (e.g., Center for Investigative Reporting, 2009). For example, the Seattle

---

1. Law & Order and the International Association of Chiefs of Police's Police Chief have for a couple of years carried prominent commercial CED advertisements featuring impressive claims about reductions in police and suspect injuries, and reducing liability risks associated with the use of force, as compared with other common less lethal technologies or techniques.

2. A search for "TASER" at YouTube (Google, Inc., Mountain View, CA) provides numerous examples from field encounters and from police training sessions.

Police Department[3] reported using CEDs in 956 incidents between January 2001 and June 2006, although we do not know how many incidents resulted in media reports (Seattle Police Department, 2006). By the end of 2008, Seattle officers had used CEDs in 1,500 instances (Seattle Police Department, 2009). Although it is possible that some incidents were associated with suspect fatalities, none are mentioned in their annual reports. As White and Ready find with their data, the nature of media interest in police use of CEDs is far more likely to reflect incidents associated with fatal outcomes than the bulk, which are less remarkable.

Media Interest and the Force Outcomes Paradox

Deadly force used by police normally is a synonym for firearm use, which mostly involves handgun discharges. From a probabilistic perspective, and somewhat paradoxically, handgun discharges by police during field encounters most likely are not going to result in fatalities. Several well-established reasons for this point are discussed in the police literature, but a key reason is that bullet hit rates in field encounters generally are well below 50%. Most of the individuals who are struck by police handgun bullets do not succumb to their wounds. The ratio of nonfatal-to-fatal outcomes has been around 2-to-1 since the 1970s, because most handgun wounds are superficial or else do not significantly damage critical organs or vessels. Prompt life-saving action taken by police at the scene also is a factor, as is the timely arrival of medical first-responders and access to emergency rooms staffed with trauma surgeons. Police handgun use, therefore, usually is not lethal, and those incidents that do not produce hits seem far less likely to result in media reports.[4]

In contrast, police use of CEDs draws media attention like a magnet when a fatality occurs. Several reasons exist for this trend, with a key one being that the term *less lethal* is too often believed to mean *nonlethal*. Another reason is that CEDs comprise the newest, widely used less lethal technology by police. This point alone makes CEDs newsworthy. CEDs deliver an electrical "shock" that might be irrationally feared or loathed by some observers because of their understanding, perceptions of, or personal experiences with electricity used in this form. Another factor that draws media attention is the still counter-intuitive instant incapacitation that CEDs often produce when used in the probe mode and the equally surprising fact that suspects nearly always immediately recover (ideally after being brought under physical control and then mechanically restrained; e.g., consider Bozeman, Hauda, Heck, Graham, Martin, and Winslow, 2009). This effect contrasts with other common force options that officers have experienced complete or partial failures with such as compliance holds, strikes and punches, baton strikes,

---

3.   I have chosen the Seattle Police Department for two reasons. White and Ready (2009) cite this department in their article, and the department maintains open access to CED-related documents on its Web site, which include annual updates from 2005 to 2009.

4.   I am not suggesting that other serious physical consequences—such as paralysis, permanent disability to a joint, or chronic pain—are not serious because they certainly are, but the focus here is on fatal outcomes.

chemical irritants, and even handguns. Such failures have been a precipitating factor in many unplanned instances of grappling and physically forcing the person to the ground.

When the use of a CED is associated with a fatal outcome, however, its improbability ensures media interest much like other phenomena that, under typical conditions, rarely result in such tragically lethal outcomes (e.g., space shuttle missions, commercial aviation, and automobile accidents). This tendency toward unusual, bizarre, or even entertaining police–public encounters naturally affects the media reports ultimately available for White and Ready (2009) to collect and analyze.

CED Policy and Training

As White and Ready (2009) note, the TASER has become the "preferred less-lethal alternative" force option among U.S. police. Blunt instruments, such as nightsticks and expandable batons, and even chemical irritants like "pepper spray" (which also experienced intense scrutiny), seem destined for reduced use. The operational potential and actual uses of CED technology is unique in police history because it offers intriguing versatility while raising questions about its legal, practical, and reasonable use.

Operationally, CED use is known to span from the lowest levels of resistance, such as forcing a passive resister to comply with lawful commands, to a substitute for deadly force in incidents in which suspects possess, brandish, or verbally threaten to use a potentially lethal weapon. Criticism has followed when police use CEDs on individuals who merely verbally or passively resist them; yet 25% of the media reports fell into this category and, interestingly, were not predictive of the fatal media reports. Although inappropriate, these uses do not seem dangerous.

Public criticism is unlikely to result from encounters at the other end of the spectrum in which, arguably, it is inappropriate to use CEDs because deadly force is an objectively reasonable option. This issue might be problematic from several perspectives, but it remains a critical matter given the restraint[5] that is often exercised by police when they are confronted with what seem to be seriously dangerous persons (e.g., see Klinger, 2004). Some early guidelines called for CED use only in nondeadly force encounters. This guideline is tricky, however, because police officers sometimes must confront individuals whose true intentions are almost impossible to deduce. In White and Ready's (2009) study, for example, 35 (6.7%) media reports involved suspects armed with guns and 79 (15.2%) involved suspects armed with knives or other cutting instruments.

Using too little force too late can have serious ramifications for officer safety, suspect injuries, and public safety; yet Klinger's (2004) extensive interviews with police officers suggest that there is a place for CEDs in some deadly force encounters. Police must be cautious, however, to avoid a 21st-century version of the classic 20th-century "Western" movie that depicts

5. By the mid-1980s, Geller (1985) had contrasted commendable and noncommendable restraint within the context of police use of deadly force.

a lawman facing down a desperado but allows the villain the first move or shot. That story line and imagery makes for great entertainment only because we can be relatively confident how the show-down will turn out. In a real police–public encounter in which death or serious bodily injury for police or others are possible, officers who feasibly could use a CED should be extremely cautious not to extend an advantage to their adversaries that could increase the danger to themselves or to the public.

Some interesting training issues revolve around the inherent accuracy of the probes as the primary CED mode, officer marksmanship ability, the single-shot design, and training cartridge cost. Because the two probes travel on paths that diverge by eight degrees, each foot of travel increases the vertical distance between them and limits the weapon's effective range (e.g., see Mesloh, Henych, Thompson, and Wolf, 2008). If the probes strike the target relatively far apart—for instance, one penetrates high in the torso and the other one low in the torso—then the current that travels between them affects a large area of muscle tissue. This electric current disrupts the normal function of a relatively greater number of motor neurons than if the probes had landed closer together. This point is important because these neurons comprise the *real* target to force incapacitation, which is distinct from voluntary compliance. Therefore, as with handgun bullets, it is of consequence *where* the CED probes strike as well as the distance between them. This issue might have implications for repeated applications of a CED during encounters that White and Ready (2009) find to be associated with fatal incidents.

Attempting to maximize the electro-muscular disruption (EMD) effect runs the risk of exceeding the practical range of the probes as projectiles. The maximum distance for the X26 might be 25 feet of wire, but research has shown that the miss rate for one probe—it does not matter which one because both are needed to complete the circuit upon impact—increases to 78% at 20 feet under laboratory conditions (National Institute of Justice, 2009a). Given the marksmanship required for effective probe firing, it is worth recalling that bullet hit rates reported in the deadly force literature actually overstate combat shooting accuracy, because they include marginal hits (e.g., an initially unnoticed, grazing bullet wound to the calf is, technically, still a hit). Police marksmanship with CEDs is important because academic researchers and some departments throughout the years have reported low levels of marksmanship with handguns, which they devote far more training to than CEDs (e.g., see Alpert, 1989; Binder and Fridell, 1984; Blumberg, 1983; Dallas Police Department, 1992, as cited in Geller and Scott, 1992; Fyfe, 1978; Geller and Karales, 1981; Kindrick, 1992, as cited in Geller and Scott, 1992; Milton, Halleck, Lardner, and Albrecht, 1977; Schade, Bruns, and Morrison, 1989). For example, bullet hit rates range from somewhat less than 1 in 4 to perhaps as high as 1 in 2 for some departments. Many of those hits, however, are superficial or otherwise incapable of forcing incapacitation upon a motivated and determined suspect. As for CEDs, we know little about hit-to-miss ratios or the influence of the circumstances at the scene, the location of officer(s) and suspect(s) relative to one another, movement on the part of one or both, and firing distances.

CEDs essentially are single-shot weapons.[6] An officer can remove a spent or faulty cartridge and replace it with a cartridge from the bottom of the frame (using an optional attachment), a pouch or pocket. This procedure takes several seconds, however, and will involve a technique that is unlikely to be as refined as speedy pistol magazine changes that are learned and practiced extensively as recruits, and thereafter maintained during in-service training or requalifying (Consider Morrison and Vila, 1998). Firing a second CED cartridge also assumes that another one is immediately available, although we do not know to what degree this availability is normal practice. In addition, suspects often are moving, and tactically, the officers often should be too, all of which compounds the CED marksmanship challenge beyond the first deployment attempt.

Because CEDs normally should not be the first choice for officers confronted with a serious deadly threat, the circumstances in which police use CEDs are less likely to be imminently dire, the tactical situation more accommodating of several responses, and the marksmanship challenge relatively less demanding. Nevertheless, in-service training and certification remain a highly discretionary area among departments. This point should come as no surprise given what we know about contemporary deadly force policies and practices in regard to instructor staffing and development, training quantity and emphases (e.g., rote requalifying versus experiential learning through several training activities), and integration with other force options through tactical decision-making scenarios (Morrison, 2003, 2006a, 2008).

Finally, "live-fire" CED training is not inexpensive; in fact, it is incredibly expensive on a shot-for-shot basis compared with the cost of conventional ammunition for handgun training. The TASER International X-26 training cartridge costs roughly $20. For the three most common service handgun calibers, those dollars would purchase approximately seventy-five 9-mm cartridges, fifty .40 Smith & Wesson cartridges, or forty .45 ACP cartridges. Live-fire CED practice, therefore, probably will be minimal and not simply because of the current national economy that already has reduced firearm training for many departments (e.g., see *American Police Beat*, 2007; Bane, 2007; Fairburn, 2007; Morrison, 2008; *Police & Security News*, 2008; also consider "The duty to train" memo, 2009). Although some departments own, or have access to, computer-based, projected-image simulation equipment, most do not. Even when they do have access, the number of scenarios that an officer would experience typically is far more limited than one might imagine. The degree to which CEDs are included in this vital training domain is another area we know little about.

---

6. The new TASER International X3 will allow for three consecutive shots in a short time frame, but given its greater bulk and the already widespread adoption of the X26, the X3 seems unlikely to be used so routinely. It might be an excellent CED, for example, for supervisors to carry in their cars for lengthy encounters, multiple suspects, and in the event of an evasive suspect in which officers need several quick shots to make the one necessary hit.

Influencing Future Practices

White and Ready (2009) note that both the Police Executive Research Forum (PERF) and the International Association of Chiefs of Police (IACP) make available model policy guidelines on the use of CEDs (PERF, 2005; IACP, 2009a, see 2009b for a larger assortment of policies). Interestingly, these guidelines carry 2005 and 2006 dates, respectively, which suggest that we already have learned much about the use of CEDs and that experiences from the past 3 to 4 years have not changed their views substantively.[7] It is not clear from this article (2009), however, how these two organizations arrived at their guidelines or whether the data continue to be grounded[8] (also consider Ederheimer and Fridell, 2005). Neither should be taken lightly, but several other important organizations have much interest in this matter and whose executives, boards, staffs, and members have various types of expertise and experience with CEDs. I briefly outline these organizations below, but all share two things in common. First, they are well positioned in their respective spheres to encourage departments to move toward model policies, training program content, delivery, and outcomes assessment that measurably improve both officer and public safety. Second, each has its limits in regard to influence.

The Commission on Accreditation for Law Enforcement Agencies (CALEA) immediately comes to mind given its mission "to improve delivery of law enforcement service by offering a body of standards, developed by law enforcement practitioners, covering a broad range of up-to-date law enforcement topics" (CALEA, 2009). Accreditation is voluntary, however, and some states now offer this form of accreditation through approved affiliates. CALEA directly impacts the roughly 900 departments that have completed the accreditation process. Its standards provide considerable latitude in the content and quantity of in-service weapon training, although these allowances can be tightened as the results from outcomes-based evaluation of field encounters become available (a topic to which I will return).

At the state level, we find police officer standards and training (POST) commissions, boards, or councils that are charged with developing, providing, and/or overseeing the delivery of comprehensive curricula for newly hired recruits. CED training, however, is not yet a common component to state-authorized or -conducted academy programs for officers from medium and smaller departments. Indeed, Reaves (2009) makes no specific mention of this type of training in his Bureau of Justice Statistics report on state and local law enforcement training academies. This oversight might change if CEDs become prevalent among a state's local departments to justify taking on this responsibility and adding some hours to the program. This addition could give POST commissions persuasive sway in shaping department in-service training; recruit programs generally comprise the most comprehensive and extensive of an officer's career, they establish minimum expectations about essential job tasks, and the future certification of CED

---

7. The IACP link to the Less-Lethal.org contains several national and international perspectives on the general use of force, which includes the use of CEDs.

8. PERF's document indicates that an extensive process was used for arriving at its 52 policy and training guidelines for police consideration (Police Executive Research Forum, 2005).

instructors could be conducted through POST courses (although in regard to handgun training, consider Morrison, 2006b).

Similarly, the International Association of Directors of Law Enforcement Standards and Training (IADLEST) influences the direction of U.S. training. For example, IADLEST's Model Minimum Standards offer guidance on training policies and programs, which stipulate that, "curriculum requirements for basic training programs should identify a set of core competencies [through job-task analysis] required for satisfactory performance of entry-level tasks" and that police officers be "proficient in the use of weapons, the ethical and effective use of deadly and *non-deadly* force and respectful of constitutional limitations on their authority" (IADLEST, 2009, emphasis added).

Two national organizations of police trainers provide excellent forums for interested instructors, supervisors, and managers to continue to develop their knowledge and skill bases, discuss persistent and emerging problems, and explore ways to advance training. First, the International Association of Law Enforcement Firearms Instructors (IALEFI), whose name suggests a singular focus, actually includes other-force-option training and certification sessions at its annual national training conference as well as at some of its regional conferences. The second organization is the International Law Enforcement Educators and Trainers Association (ILEETA). Although its annual conference covers a wide range of topics, the full use of force spectrum is heavily emphasized. Neither IALEFI nor ILEETA, however, can mandate compliance with what they might consider to be best practices. This issue leaves these and other membership associations dependent on voluntary adoption of recommended standards.

Closing Thoughts

Police use of CEDs should be guided by empirically informed policies and training. Although dramatic successes and failures are most likely to be brought to our attention through media reports, we need to know more about the full range of outcomes. It is essential to keep in mind that rigorous investigation of CED deployments will usefully add to our knowledge and best serve practitioners as we strive to refine technologies, policies, and training. This refinement is central to the dual goal of enhancing police performance in violent encounters while maximizing officer and public safety.

White and Ready (2009) extensively discuss the weaknesses of using media reports as a data source for CED deployments and outcomes as well as their analytical results' dependency on the media report portrayals. Data collected from media reports of police use of CEDs seems to be suited best for examining media representations of these particular police–public encounters. It is difficult to ascertain how much this approach enlightens us on deployment outcomes, because the nature and extent of information conveyed by media reports is constrained by the reporters' understanding of the information they have been provided, their interpretation of the facts as they know them, their degree of familiarity with CED technology and relevant policies

and training, as well as the available column inches and editorial priorities.[9] Kobler (1975a, 1975b) faced a similar challenge in the late 1960s during the exploratory period of deadly force research. He relied on a newspaper clipping service to collect information on 1,500 fatal police shootings in the United States (see Fyfe, 1978, for a discussion of Kobler's study within the context of that era).

Today, academic and practitioner researchers have many research and program evaluation methodologies available to them for studying police use of force. Furthermore, the National Institute of Justice (NIJ) and other organizations are funding multiple projects specifically on the topic of less lethal technology, especially CEDs. Findings and analyses from some of these projects already are being presented at professional conferences and will appear in academic journals. For example, NIJ currently is funding academic and/or medical researchers in their studies of police reports, autopsies, toxicological analysis, post-CED application medical re-cords, and immediate care; the chains of events that occur prior to CED incidents associated with fatalities; the effects of CEDs on internal organs, which include the cardio and pulmonary systems; individuals who police subdued with CEDs and admitted to a hospital; and excited delirium as a contributing factor in fatal incidents (National Institute of Justice, 2009b). As the empirical database grows, POST commissions, accreditation bodies, professional associations, and the courts will have an increasingly rich foundation for influencing police CED policies and practices.

One of the most persistent shortcomings to police use-of-force training, however, has been the lack of feedback from the field specifically designed to provide rich data on outcomes to both practitioners and academicians that span the full range of encounters. This feedback is the cornerstone to maximizing police performance and, thereby, to enhancing both officer and public safety. This could produce the empirical foundation for generating compelling standards. Producing the data, analyzing it, and sorting out its implications for police use of CEDs will provide ample opportunities for interested researchers and practitioners to make groundbreaking and critically important contributions to the literature and police practices.

Implementing best practices in the form of continuously refined policies and robust train-ing programs seems to be beyond the capacity of a single entity. A working group comprised members from the governmental and nongovernmental organizations discussed earlier, however, would be indispensable in shaping, steering, and encouraging the use of new knowledge that could flow from an applied research and evaluation agenda. Fortunately, academics and police practitioners both place a high value on tangible results, and the organizations discussed earlier place officer safety at or near the top of priorities. Collaborative research between academic and practitioner researchers would bring the added benefit of lowering barriers even more to other important police research projects. A national task force comprised key stakeholders that are focused on police use of CEDs in a democratic society—one that expects high levels of officer

9.   Also worth noting is that 37.5% of the 521 nonduplicate cases were from California, Colorado, and Florida.

judgment and technical proficiency but also seeks to enhance officer and public safety—seems capable of producing meaningful and lasting contributions.

References

Alpert, Geoffrey P. 1989. *Metro-Dade Police Department Discharge of Firearm Study: 1984–1988*. Unpublished consulting report to the Metro-Dade Police Department.

Ammunition shortage affects law enforcement training. 2008. *Police & Security News*, 24: 47.

Bane, Tim. 2007. Are instructors providing adequate firearms training? *The Firearms Instructor,* 43: 5–6.

Binder, Arnold and Loire A. Fridell. 1988. Lethal force as a police response. *Criminal Justice Abstracts*, 16: 250–280.

Blumberg, Mark L. 1982. *The use of firearms by police officers: The impact of individuals, communities, and race*. Ann Arbor, MI: University Microfilms International.

Bozeman, William P., William E. Hauda II, Joseph J. Heck, Derrel D. Graham, Brian P. Martin, and James E. Winslow. 2009. Safety and injury profile of conducted electrical weapons used by law enforcement officers against criminal suspects. *Annals of Emergency Medicine,* 53: 480–489.

Center for Investigative Reporting. 2009. *Latest statistics*. Retrieved August 27, 2009 from centerforinvestigativereporting.org/files/LatestStats.pdf.

Commission on Accreditation of Law Enforcement Agencies, Inc. 2009. About CALEA. Retrieved August 27, 2009 from calea.org/Online/AboutCALEA/Commission.htm.

Dallas Police Department. 1992. *Comparative statistical survey: 1991*. Unpublished report.

Fairburn, Dick. 2007. Surviving the ammo shortage. *The Police Marksman,* 32: 30–32.

Ederheimer, Josh A. and Lorie A. Fridell. 2005. *Chief concerns: Exploring the challenges of police use of force*. Washington, DC: Police Executive Research Forum.

Fyfe, James J. 1978. *Shots fired: An examination of New York City Police firearms discharges*. Ann Arbor, MI: University Microfilms International.

Geller, William A. 1985. Officer restraint in the use of deadly force: The next frontier in police shooting research. *Journal of Police Science and Administration*, 13: 153–171.

Geller, William A. and Kevin J. Karales. 1981. *Split-second decisions: Shootings of and by Chicago police*. Chicago, IL: Chicago Law Enforcement Study Group.

Geller, William A. and Michael S. Scott. 1992. *Deadly force: What we know*. Washington, DC: Police Executive Research Forum.

International Association of Chiefs of Police. 2009a. *Developed model policies*. Retrieved August 27, 2009 from less-lethal.org/web/articles.aspx?group=4.

International Association of Chiefs of Police. 2009b. *Law-enforcement agency policies*. Retrieved August 27, 2009 from less-lethal.org/web/articles.aspx?group=2.

International Association of Directors of Law Enforcement Standards and Training. 2009. *Model minimum standards*. Retrieved August 27, 2009 from iadlest.org/.

International Association of Law Enforcement Firearms Instructors. 2008. *About IALEFI*. Retrieved on February 11, 2008 from ialefi.com/mission.htm.

Kindrick, Roy E. 1992. *Officer-involved deadly force incident report.* Unpublished Portland, OR, Police Bureau Training Division memorandum.

Klinger, David. 2004. *Into the kill zone.* San Francisco, CA: Jossey-Bass.

Kobler, Arthur L. 1975a. Police homicide in a democracy. *Journal of Social Issues,* 31: 163–184.

Kobler, Arthur L. 1975b. Figures (and perhaps some facts) on police killing of civilians in the United State, 1965–1969. *Journal of Social Issues,* 31: 185–191.

Mesloh, Charlie, Mark Henych, L. Frank Thompson, and Ross Wolf. 2008. *A qualitative & quantitative analysis of conducted energy devices: TASER X26 vs. Stinger S2000.* Unpublished report. National Institute of Justice. Retrieved on August 27, 2009 from ncjrs.gov/pdffiles1/nij/grants/222769.pdf.

Milton, Catherine H., Jenne W. Halleck, James Lerner, and Gary L. Albrecht. 1977. *Police use of deadly force.* New York: Wiley.

Morrison, Gregory B. 2003. Police and correctional department firearm training frameworks in Washington State. *Police Quarterly*, 6: 192–221.

Morrison, Gregory B. 2006a. Deadly force programs among larger U.S. police departments. *Police Quarterly*, 9: 331–361.

Morrison, Gregory B. 2006b. Police department and instructor perspectives on pre-service firearm and deadly force training. *Policing: An International Journal of Police Strategies & Management*, 29: 226–245.

Morrison, Gregory B. 2008. *Police firearms training survey, 2008: Preliminary findings on policies, programs and practices.* Paper presented at the annual meetings of the American Society of Criminology, St. Louis, Missouri.

Morrison, Gregory B. and Bryan J. Vila. 1998. Police handgun qualification: Practical measure or aimless activity? *Policing: An International Journal of Police Strategies and Management*, 21: 510–533.

National Institute of Justice. 2009a. *Comparing the TASER X26 and the Stinger S2000.* Retrieved on August 27, 2009 from ojp.usdoj.gov/nij/topics/technology/less-lethal/TASER-stinger.htm.

National Institute of Justice. 2009b. *Assessing the safety of conducted energy devices.* Retrieved August 27, 2009 from ojp.usdoj.gov/nij/topics/technology/less-lethal/conducted-energy-devices.htm.

Police Executive Research Forum. 2005. *PERF conducted energy device policy and training guidelines for consideration.* Retrieved August 27, 2009 from policeforum.org/upload/PERF-CED-Guidelines-Updated-10-25-05%5B1%5D_715866088_1230200514040.pdf.

Police waiting up to one year for ammo: In high demand and short supply. 2007. *American Police Beat*, 14: 4.

Reaves, Brian A. 2009. *State and local law-enforcement training academies, 2006.* U.S. Department of Justice, Bureau of Justice Statistics. Retrieved August 27, 2009 from ojp.usdoj.gov/bjs/pub/pdf/slleta06.pdf.

Schade, Thomas V., Gilbert H. Bruns, and Gregory B. Morrison. 1989. Armed confrontations: Police shooting performance in armed confrontations. *American Journal of Police,* 8: 31–48.

Seattle Police Department. 2006. *TASER use update.* Retrieved August 27, 2009 from seattle. gov/police/publications/Special/TASERUseUpdate0706.pdf.

Seattle Police Department. 2009. *TASER use update.* Retrieved August 27, 2009 from seattle. gov/police/publications/Special/TASERUseUpdate0509.pdf.

TASER International, Inc. 2009. *TASER X26.* Retrieved August 27, 2009 from TASER. com/products/law/Pages/TASERX26.aspx.

The duty to train officers is unaffected by the lack of reimbursement sources. 2009. 2003 memorandum republished by California Police Chiefs Association, Sacramento.

White, Michael D. and Justin Ready. 2009. Examining fatal and nonfatal incidents involving the TASER: Identifying predictors of suspect death reported in the media. *Criminology & Public Policy.* This issue.

Gregory B. Morrison earned his Ph.D. at the University of California at Irvine and joined the Ball State University faculty in 1999. He has been a municipal police officer, trainer, and manager. His primary teaching interests are U.S. policing together with research methods and data analysis. Currently, Morrison is researching U.S. police department policies, programs, and practices that revolve around firearm training, assessment of skills and abilities, and the use of deadly force.