❈ UNDERSTANDING POLICE GANG DATA AT THE AGGREGATE LEVEL: AN EXAMINATION OF THE RELIABILITY OF NATIONAL YOUTH GANG SURVEY DATA (RESEARCH NOTE)

Charles M. Katz
Center for Violence Prevention and Community Safety
School of Criminology and Criminal Justice
Arizona State University

Andrew M. Fox
Department of Criminal Justice and Criminology
University of Missouri- Kansas City

Chester L. Britt
School of Criminology and Criminal Justice
Northeastern University

Phillip Stevenson
Arizona Criminal Justice Commission

❈ Abstract

The National Youth Gang Survey (NYGS) is an annual survey of a nationally representative sample of police agencies administered by the National Gang Center that provides data on the number of gangs, gang members, and gang homicides across jurisdictions. The focus of the current study is the reliability of key measures of gang behavior measured by the NYGS. Specifically, the current study assesses NYGS data from 2005 to 2009 in terms of missing data, test-retest reliability, internal consistency reliability, and inter-observer reliability. Results showed low levels of missing data for number of gangs and number of gang homicides in jurisdictions with fewer than 200,000 people, while measures of number of gang members and gang homicides in jurisdictions over 200,000 people had missing data from about 30% of participating jurisdictions. The findings also suggest that the NYGS data on gangs and gang members are generally highly reliable. Data on gang homicides also exhibited a high degree of reliability, but only for jurisdictions with populations over 200,000. Implications of the findings and recommendations for future research are discussed.

Macro-level gang research has been sparse when compared to individual-level research because of the lack of valid and reliable data (Klein & Maxson, 2006). The macro-level data available on gangs, gang members, and gang crimes are largely the product of organizational surveys of police agencies. The surveys typically contain questions regarding the number of gangs, gang members, and amount of gang crime within the agency's jurisdiction in a given year (Curry, Ball, Fox, & Stone, 1992; Miller, 1982; Egley, Howell, & Moore, 2010). While these data have been available for more than a decade, some have questioned their utility for answering questions regarding theory, policy, and practice (Zatz, 1987; McCorkle & Miethe, 1998; Katz, 2003), in part because research on the validity of police-generated gang data at the agency level has been mixed (e.g., Meehan, 2000; Zatz, 1987; McCorkle & Miethe, 1998). Early reports were extremely negative, noting the many institutional, political, and human factors that affected these data's validity (Hagedorn, 1990; Zatz, 1987; Katz, 2003; Webb & Katz, 2003, 2006). However, more recent work has shown that the data accurately discriminate between those who are a greater threat to a community and those who are not (Katz, Webb, & Schaefer, 2000). Unfortunately the research published on police gang data so far has been largely limited to single-city accounts (Chesney-Lind, Rockwell, Marker, & Reyes, 1994; Zatz, 1987; McCorkle & Miethe, 1998; Katz et al., 2000; Katz & Webb, 2006; for exception see Decker & Pyrooz, 2010).

Perhaps the most widely available data on gangs, gang members, and gang crimes nationally come from the National Youth Gang Survey (NYGS). The National Gang Center is funded by the Office of Juvenile Justice and Delinquency Prevention (OJJDP) to collect official gang data through the NYGS from a nationally representative sample of law enforcement agencies (National Youth Gang Center, 2008).[1] The average annual response rate is about 85%, and many of the same agencies have been surveyed continuously since 1996. This allows the data to be used for the longitudinal analysis of gang problems in many jurisdictions across the nation (National Youth Gang Center, 2009). The problem, however, is that the reliability of the NYGS data is still relatively unknown. Only one study to date has systematically examined them, and it was limited to homicide data in the largest 100 jurisdictions (Decker & Pyrooz, 2010). The present study seeks to expand on this inquiry by examining data derived from the NYGS in four ways: 1) the extent of missing data, 2) test-retest reliability, 3) internal consistency reliability, and 4) inter-observer reliability.

———————————————

[1] The name of the National Youth Gang Center was recently changed to the National Gang Center. For the purpose of this paper we use the term National Youth Gang Center as well as National Gang Center depending on the year the data were collected or report was published.

❈ Literature Review

Collecting information systematically about gangs is a challenge for several reasons. While there are a number of studies that use self-report surveys to estimate levels of gang involvement (see Klein & Maxson, 2006, p. 23), many of them are cross-sectional and location specific, thus not allowing for longitudinal assessment or estimates at the national or city level. An alternative method for determining the prevalence of gangs, gang members, and gang crime is to survey law enforcement agencies about their knowledge of these issues (Klein, 1995). However, a number of concerns have been raised about the validity of such data.

Concerns about the validity and reliability of police gang data can generally be classified into two major categories. One concern is that police gang data are unreliable as a result of police administrative practices, specifically, problems associated with administrative definitions (Maxson & Klein, 1990; Klein 1995) and lack of quality control (Katz, 2003; Hagedorn, 1990). The debate over the definition of gang, gang members, and gang crime has been long-standing (see Klein & Maxson, 2006). Researchers and policymakers alike have relied on several methods over the years to identify gang members and gang crimes. However, because of the many definitions that have been created and institutionalized (Barrows & Huff, 2009), several scholars questioned their subsequent impact on the measurement of gang-related phenomena (Maxson & Klein, 1990, 1996). For example, Maxson and Klein (1990) applied two definitional approaches to homicides recorded in Los Angeles. One definition required evidence of motivation for a homicide to be recorded as a gang homicide, while the other definition recorded the homicide as a gang homicide if either the offender or victim was a gang member. When the authors applied the more restrictive motive-based definition to the homicide data, the estimated rate of gang homicides was reduced by approximately 50% compared to the broader member-based definition. Maxson and Klein concluded that definitional inconsistencies between departments make it difficult to generalize about the magnitude of a gang problem from one community to another. Other critics have argued that even if police agencies used a common definition they would not possess the quality control mechanisms to ensure that definitions are being used effectively and equitably. For example, Katz (2003) examined how the construction of gang statistics in one Midwestern agency was influenced by serious abnormalities in internal information processing. Gang statistics related to gangs, gang members, or gang homicides were not found to be the product of the application of official definitions, or even informal definitions, but rather the product of inadequate communication both within the gang unit and between the gang unit and its operating environment.

A second concern about police gang data is that they are highly susceptible to manipulation. Some academics have reported through case studies that law enforcement agencies purposefully distorted gang statistics to fulfill organizational needs (Hagedorn, 1990), obtain resources (Zatz,1987), or repair the image of their

agency (McCorkle & Miethe, 1998). Others have voiced concern over the fact that police gang data are subject to uncontrolled officer discretion (Burrell, 1990; Crew, 1997), which can present an inaccurate view of the gang problem. In Denver and Los Angeles, for example, it was reported that about two thirds of African-American youth were documented by the police as gang members (Crew, 1997).

At least three studies to date have examined the external validity of police gang data. In somewhat similar studies, Zatz (1985, 1987) and McCorkle and Miethe (1998) used court data to measure the objective threat of gangs at the community level. Zatz (1985, 1987) examined juvenile court records in Phoenix, Arizona, comparing documented gang members to a matched sample. She reported that when compared to non-gang members, documented gang members were no more likely to engage in violent crime or serious narcotics offenses. Zatz (1987) concluded by noting that the increasing number of documented gang members in Phoenix was a consequence of the police department's need to demonstrate a growing gang threat in order to obtain access to federal grant dollars to establish a gang squad. Likewise, McCorkle and Miethe (1998) studied the threat posed by gangs in Las Vegas, Nevada, by using adult court records to examine the proportion of index crimes and felony drug charges filed against documented gang members. They argued that their analysis suggested that "The actual threat posed by gangs…was less real than imagined" (p. 58) and that Las Vegas' purported problem could be attributed to a moral panic constructed by the police department to obtain resources and to regain legitimacy within their community. The authors concluded that the creation of the moral panic was made easy since the agency simply documented individuals who were either on the fringe of gang membership or who wore gang-style clothing. Similarly, research conducted in Hawaii by Chesney-Lind and colleagues (1994) examined the differences between gang and non-gang youth in terms of criminal involvement. While their sample consisted of a small number of gang members (n = 78), the authors similarly concluded that "whatever the designation of youth gang used by the [Honolulu Police Department], it does not denote delinquent/criminal behavior that substantially differentiated these youth" (Chesney-Lind et al., 1994, p. 218).

Katz (2003) examined the production and dissemination of gang statistics in one large Midwestern police gang unit. He found major problems associated with compilation of the agency's gang data. Much of this, he reported, was the result of administrative policies and procedures being unknown by those who documented gangs, gang members, and gang crimes. He further noted that gang-related phenomena were not documented according to departmental policy but rather in an ad hoc, irregular manner. The author's observations led him to question the reliability and validity of the agency's gang data.

One of the few studies to demonstrate the discriminate validity of police gang data was conducted by Katz, Webb, and Schaefer (2000) in Mesa, Arizona. The authors compared 373 police identified gang members to 1,832 non-gang members and found that gang members were significantly more likely to be arrested for almost every crime type (e.g. violent, weapons related, property, drug, status).

Further, the authors found that distinctions made by the police between gang members and gang associates showed further promise of separating the more criminally involved from those who were marginally involved. The study's findings are in contrast to what earlier research suggested. In summary, the authors found that "the police department was able to identify and document youth that pose a more serious threat to the community" (Katz et al., 2000, p. 431). Thus, the literature on the validity of official gang data is somewhat mixed.

⊗ National Youth Gang Survey Data

The National Gang Center (NGC) was established in 1995 by the Office of Juvenile Justice and Delinquency Prevention. Since 1996, the NGC has conducted the National Youth Gang Survey (NYGS). The NYGS is conducted annually and includes a nationally representative sample of law enforcement agencies from across the United States. The survey asks law enforcement agencies to report the extent and nature of gang membership in their jurisdiction. The methodology includes the following sample: 1) all police departments serving cities with populations of 50,000 or more; 2) all suburban county police and sheriffs' departments; 3) a randomly selected sample of police departments serving cities with populations between 2,500 and 49,999; and 4) a randomly selected sample of rural county police and sheriffs' departments. The NYGS defines a youth gang as: "a group of youths or young adults in your jurisdiction that you or other responsible persons in your agency or community are willing to identify or classify as a 'gang'" (National Youth Gang Center, 2009). Additionally, the survey instructions asked respondents not to include motorcycle gangs, hate or ideology groups, prison gangs, or other exclusively adult gangs.

It is important to examine the properties pertaining to the reliability of the NYGS for a number of reasons. First, researchers and policymakers need a reliable method of measuring the current scope of the gang problem (Barrows & Huff, 2009; Kennedy, 2009; Klein, 2009; Short, 2009). The NYGS offers valuable trend information that can be useful at the local and national levels. Second, the NYGS, if reliable, offers a rich source of information to evaluate programs and strategies to combat gangs and gang violence (Klein, 2009). Third, the information collected as part of the NYGS can assist in the allocation of financial resources that are concerned with the gang problem. And fourth, this type of data can be useful in furthering theoretical explanations of gang formation. The availability of a reliable, nationally representative database on gangs would be valuable and is needed for future academic research (Katz & Schnebly, 2011). Further, these data can be combined with other national data sets to produce informative research (i.e., U.S. Census, LEMAS, etc.; Katz, Maguire, & Roncek, 2002; Pyrooz, Fox, and Decker, 2010).

To our knowledge only one study has examined the quality of these data. It was conducted by Decker and Pyrooz (2010), who examined data from the largest 100 cities in the United States (i.e., cities with a population of 200,000 persons or

more) from 2002 through 2006. The authors examined the validity and reliability of gang homicide data obtained from Uniform Crime Reports (UCR), Supplementary Homicide Reports (SHR), and the NYGS. The authors reported that while SHR data were insufficient for analysis of gang homicides, NYGS was a reliable and valid source of data on gang homicides. They argued that based on their results, researchers and policymakers should make greater use of these data for both cross-sectional and longitudinal research.

While their study was a major advance in researchers' understanding of the reliability and validity of large-city gang homicide data, it was limited in several important ways. First, their study did not examine the validity and reliability of NYGS measures of gangs or gang members. As discussed above, such data have important theoretical and policy implications. Second, the study did not examine variation in the reliability of the measures of gang homicide by region, population size, or issues related to missing data. Prior research examining official police data suggests that variation in reliability and missing data does not occur randomly but is often related to these factors (Maguire & Katz, 2002; Fox, 2004). Third, Decker and Pyrooz' (2010) analysis examined the reliability and validity of NYGS homicide data as compared to UCR and SHR data. These two data sources serve as poor lenses with which to examine the reliability and validity of gang homicide data because of the vast difference in measurement. UCR homicide data capture information on all homicides, not just gang homicides, which vary in their contribution to community levels of homicide (Klein & Maxson, 2006). SHR data is only collected for juvenile gang homicides, and, as Maxson (1999) has pointed out, most gang-involved homicides involve adults and not juveniles.

⊗ The Present Study

As noted above, the body of literature examining the reliability of police-driven gang data suffers from some notable limitations. First, the above research is limited to either single-city case studies (Chesney-Lind et al, 1994; Zatz, 1987; McCorkle & Miethe, 1998; Katz et al, 2000; Katz, 2003) or duel-city examinations of the impact of definitions on gang crime (Maxson & Klein, 1990; 1996). Second, the only study that examined the validity and reliability of national youth gang survey data, Decker and Pyrooz, 2010, was limited in several ways, including the restriction of the analysis to gang homicide data from large cities.[2] The current study builds upon Decker and Pyrooz' work by examining the reliability of police gang data

---

[2] For an exception, see Pyrooz et al., 2010, who conducted basic analyses to examine the validity of these data, and found them to be sufficiently valid for use in their analysis of the relationship between macro-level factors and gang membership (see p. 875, paragraph 2).

from multiple cities. Specifically, the present study examines whether the NYGS is a source of reliable data in four ways.[3]

First, we examine the extent of missing data to determine their availability to researchers and policymakers and to determine if missing data occur randomly. Substantial amounts of missing data can result in researchers and policymakers modifying their analysis by reducing their sample size, which in turn can reduce statistical power. Furthermore, if there is a substantial amount of missing data, and the missing data differ significantly from the non-missing data (i.e., systematically missing data), findings can be biased (Allison, 2001; Horton & Lipsitz, 2001).

Second, we examine test-retest reliability. Test-retest reliability measures the consistency of the measurement from one point in time to another. While we expect the number of gangs, gang members, and gang homicides to change from year to year, these changes should be subtle and not drastic. If the number of gangs, gang members, or gang homicides changes significantly and frequently in a large number of jurisdictions over the five-year study period, it might suggest that the data do not accurately reflect the desired constructs. Thus, we expect that there should be a high level of consistency between years if the measures are reliable (Maxfield & Babbie, 2011).

Third, we examine internal consistency reliability. Internal consistency reliability measures the consistency between items at the same measurement point. Prior research has measured community gang problems generally through the number of gangs, gang members, and gang homicides. These measures have often been used interchangeably, resulting in researchers and policymakers implicitly suggesting that the three measures tap into the same underlying construct. Therefore, if the numbers of gangs, gang members, and gang homicides measure the same underlying construct, we would expect, for example, that those communities that have a large number of gangs also have a large number of gang members and gang homicides, and those with a large number of gang members would have a large number of gangs and gang homicides (Trochim, 2006). If there is not internal consistency between these measures it might suggest that the data are either not reliable or that the measures do not measure the same underlying construct.

Last, we examine inter-observer reliability. Inter-observer reliability (also called inter-rater reliability) is the extent to which different sources or observers consistently measure the same phenomena. We know little about the inter-observer reliability of police organizational surveys, and the little research that has been conducted suggests that large-scale police organizational surveys such as the Law

---

[3] Validity refers to the accuracy of a measure and whether the variable reflects what the researcher is trying to measure. Reliability refers to the consistency of the measure and whether the researcher will obtain the same information after repeated measurement (Weisburd & Britt, 2003; Maxfield & Babbie, 2011). For further discussion of the issues see Katz, Webb, Gartin, & Marshall, 1997, and Webb, Katz, & Decker, 2006.

Enforcement Management and Administrative Statistics (LEMAS) survey may lack inter-observer reliability (Walker & Katz, 2005). In the present study we are interested in whether there is inter-observer reliability with NYGS data. We expect that if NYGS data are reliable, organizations would generally report the same information on two similar surveys administered at roughly the same time.

⊠   Methods

The data analyzed in this article were collected by the National Gang Center from 2005 through 2009. Since 1996, the National Youth Gang Center has been under contract from the Office of Juvenile Justice and Delinquency Prevention to collect gang-related data for the National Youth Gang Survey. The NYGS is administered annually to a nationally representative sample of law enforcement agencies, with an average response rate of about 85%. The survey instrument includes questions that focus on the number of gangs, gang members, and gang homicides in a community, as well as questions that focus on the demographic characteristics of gangs and gang members, and their community's responses to gangs and gang problems. For the present study we restricted our analyses to agencies that serve populations of 25,000 or more persons.[4]

We examine the inter-observer reliability of the NYGS data by comparing them to data obtained through the Arizona Gang Threat Assessment. Since 1990, the Arizona Criminal Justice Commission (ACJC) has been required by state statute to conduct a statewide survey of criminal justice agencies to assess the scope and nature of the state's gang problem. From 2007 through 2009, ACJC administered a survey to all federal, state, county, municipal, tribal, university, and airport law enforcement agencies in Arizona. For the purposes of this study, we only examine the data received from the municipal police agencies (n = 12) and sheriffs' departments (n = 6) in the state that could be matched to those agencies participating in the NYGS. Response rates for the Arizona Street Gang Survey vary by year and ranged between 72% and 88%. The Arizona Gang Threat Assessment included two similar questions that were asked by the National Youth Gang Survey. These questions pertained to the presence of gangs and the number of gang members in the jurisdiction, which are the subject of the present study.

Measures

The present study examines four variables from the NYGS that are of particular interest to researchers and policymakers. The first variable of interest is the filter

---

[4] Jurisdictions with populations under 25,000 were excluded from the study because the probability of these communities being repeatedly sampled over the five-year period and consistently reporting gang activity (and therefore providing data) was small.

question that is asked by the NYGS: "During 20XX, were any youth gangs active in the city, town, or county served by your agency?" The second is the number of gangs in the jurisdiction. The NYGS instrument asked respondents, "How many youth gangs (as defined in the Survey Instructions) were active in your jurisdiction during 20XX?" [5] The third variable of interest is the number of gang members in the jurisdiction. The NYGS instrument asked, "How many youth gang members were active in your jurisdiction during 20XX?" An active gang member was defined by the NYGS as someone who was involved in gang activity with other gang members during the given year. The fourth variable of interest was the number of gang homicides that occurred in the jurisdiction. Each responding agency was asked to provide the number of member-based homicides, including those in which "a gang member(s) is either the perpetrator or the victim, regardless of motive" (2008 National Youth Gang Survey, 2009, p. 4).

Additionally, two measures from the Arizona Gang Threat Assessment were used to examine the inter-observer reliability of the NYGS data. The first was whether gangs were present in the jurisdiction. Respondents were asked "Are there any gangs or gang members in your jurisdiction?" Second, respondents were asked about the number of gang members in their jurisdiction: "What is the total membership of gangs in your area?"

We examine the reliability of the NYGS data in the context of two environmental variables for the purpose of examining potential sources of bias. The first is the size of the population that the law enforcement agency serves. Prior research has found that the size of a community has an impact on the scope and nature of the gang problem (Klein, 1995), and is related to agency recording practices with respect to gang problems (Curry, Ball, & Fox, 1994). Based on this body of literature, we hypothesize that agencies located in larger jurisdictions provide more reliable data for the National Youth Gang Survey. One reason is that agencies located in large jurisdictions are more likely to be large themselves and have gang problems (Klein & Maxson, 2006), and are thus more likely to have bureaucratic mechanisms in place for formally recording this information (Katz et al., 2002). Second, and related, agencies located in large jurisdictions are more likely to have a formalized mechanism that is responsible for collecting, maintaining, and disseminating information on gangs (Langton, 2010). For the present study, jurisdictions were divided into four categories. Those with

---

[5] The survey instructions state, "For the purpose of this survey, a 'youth gang' is defined as: a group of youths or young adults in your jurisdiction that you or other responsible persons in your agency or community are willing to identify or classify as a 'gang.' Do not include motorcycle gangs, hate or ideology groups, prison gangs, or their exclusively adult gangs. Please base your responses on your records, your personal knowledge, and/or consultation with other agency personnel who are familiar with youth gangs" (2008 National Youth Gang Survey, 2009, p. 1).

populations between: 1) 25,000 through 49,999; 2) 50,000 through 99,999; 3) 100,000 through 199,999; and 4) those with 200,000 or more.

The second environmental variable examined is region of the country. For this we used the four primary regions defined by the Federal Bureau of Investigation and the Census Bureau: West, Midwest, South, and Northeast. There are two primary reasons we chose this environmental variable. The first is that prior research suggests that gang problems have been found to vary by region, with the West having the most serious gang problem, followed by the Midwest and South. Jurisdictions in the Northeast are far less likely to report gang problems (Klein, 1995). Second, policing scholars have repeatedly found that policy and program innovations vary by region (King, 2000), and that agencies in some regions are more innovative than others (Wilson, 1968; Langworthy & Travis, 1994). For example, Steel and Lovrich (1987) reported that police agencies located in the South and West were significantly more likely to hire female sworn officers, Maguire and Katz (2002) reported that police agencies located in the Midwest were more likely to claim to implement community policing, and Hassell and Zhao (2003) reported that police agencies in the Northeast allow for greater discretion among frontline supervisors in the era of community policing.

Analytic Strategy

We first examine the filter question that is asked by the National Youth Gang Survey, "During 20XX, were any youth gangs active in the city, town, or county served by your agency?" to examine the stability of reporting a gang problem. We next examine each of our gang measures for missing data or "don't know" responses by year. Following these analyses, we examine the test-retest reliability of the gang, gang member, and gang homicide data by examining correlations between each of the five years of data collection. We then examine internal consistency reliability by assessing the consistency of results across our three measures (i.e., gang, gang member, and gang homicide) by year. Last, we examine the inter-observer reliability of our gang measures by using data from both the National Youth Gang Survey and the Arizona Gang Threat Assessment.

▓ Findings

Table 1 shows the percentage of jurisdictions that reported having active youth gangs by year. Across the five years of data the pattern is the same: the larger the city, the more likely it is that the city reported having a gang problem. For example, in 2009, 46.9% of cities with a population between 25,000 and 49,999 reported having a gang problem, compared to 74.9% of cities with a population between 50,000 and 99,999, 88.1% with a population between 100,000 and 199,999, and 96.0% of cities with a population greater than 200,000. The presence of active gangs reported by a jurisdiction also varied significantly by region of the country.

◈ Table 1

Percentage of Jurisdictions Reporting Active Youth Gangs

	2005			2006			2007			2008			2009		
	No	Yes	N	No	Yes	N	No	Yes	N	No	Yes	N	No	Yes	N
Population															
25,000 thru 49,000	56.4%	43.6%	225	52.7%	47.3%	283	55.8%	44.2%	269	55.7%	44.3%	273	53.1%	46.9%	277
50,000 thru 99,999	29.3	70.7	423	28.0	72.0	492	26.8	73.2	451	27.0	73.0	422	25.1	74.9	482
100,000 thru 199,000	9.9	90.1	172	7.8	92.2	205	10.5	89.5	190	12.1	87.9	174	11.9	88.1	201
200,000 thru high	8.2	91.8*	110	4.6	95.4*	131	1.7	98.3*	119	2.4	97.6*	123	4.0	96.0*	124
Region															
Midwest	41.2	58.8	228	45.0	55.0	271	44.1	55.9	261	44.5	55.5	254	39.9	60.1	271
Northeast	36.6	63.4	142	31.0	69.0	168	35.9	64.1	145	36.5	63.5	148	39.5	60.5	147
South	32.7	67.3	327	26.8	73.2	410	27.2	72.8	394	29.1	70.9	371	27.2	72.8	416
West	10.3	89.7*	233	9.5	90.5*	262	8.3	91.7*	229	6.8	93.2*	219	7.2	92.8*	250
Total	29.8	70.2	930	27.8	72.2	1,111	28.5	71.5	1,029	29.2	70.8	992	27.4	72.6	1,084

* p < .05; chi-square difference within year by population or region.

Across years, western jurisdictions were the most likely to report the presence of active gangs, followed by cities in the South, Northeast, and Midwest. Specifically, in 2009, 92.8% of western jurisdictions, 72.8% of southern jurisdictions, 60.5% of northeastern jurisdictions, and 60.1% of midwestern jurisdictions reported active youth gangs.

Missing Data

Table 2 displays the percentage of jurisdictions that had missing data or recorded a "don't know" response. The results in Table 2 are partitioned by population size and region of the country for each measure —number of gangs, number of gang members, and number of gang homicides. For our measure of number of gangs there were few significant differences by jurisdiction size or region. The only exception was for the year 2005, where those jurisdictions with a population of more than 200,000 residents were significantly less likely to provide data on the number of gangs in their jurisdiction. Over the five-year period examined, on average 17.2% of jurisdictions with a population between 25,000 and 49,999, 13.3% of jurisdictions with a population between 50,000 and 99,999, 11.3% of jurisdictions with a population between 100,000 and 199,999, and 15.3% jurisdictions with a population 200,000 or more did not report the number of gangs.

We found a substantial amount of missing data and "don't know" responses for the number of gang members. Specifically, about one third of jurisdictions did not provide data on the number of gang members in their jurisdiction. However, the provision of these data did not vary by population or region from 2005 through 2009.

With respect to gang homicides, missing data and "don't know" responses significantly varied by population and region across several of the study years. For all five study years the smallest jurisdictions had the least missing data and the largest cities had the most missing data. The five-year average, for example, depicted in Table 2, shows that about 30% of jurisdictions with a population of 200,000 or more did not provide data on the number of gang homicides, compared to 17.3% of jurisdictions with a population of 100,000 to 199,999, and about 12% of jurisdictions with a population of 25,000 to 99,000 residents. In 2008 and 2009, region was also significantly associated with a jurisdiction not providing data on the number of gang homicides. For both years, those jurisdictions located in the Midwest and South were significantly less likely to provide gang homicide data when compared to jurisdictions in the West and Northeast. For instance, in 2009, 10.1% of jurisdictions in the Northeast, 12.9% of jurisdictions in the West, 19% of jurisdictions in the Midwest, and 19.8% of jurisdictions in the South did not provide data on the number of gang homicides.

Test-Retest Reliability Among Measures

Table 3 displays the test-retest reliability of our measures by year. The correlation coefficients displayed in the table are the correlations between two years on

✦ Table 2

Percentage of Communities Providing Missing Data or "Don't Know" Response

	2005	2006	2007	2008	2009	5-Year Average
Number of Gangs						
Population						
25,000 thru 49,999	17.3	14.2	18.5	17.4	18.5	17.2
50,000 thru 99,999	9.4	13.6	10.6	18.8	13.9	13.3
100,000 thru 199,999	7.7	12.2	8.8	16.3	11.3	11.3
200,000 thru high	21.8*	12.0	10.3	17.5	15.1	15.3
Region						
Midwest	16.4	14.1	12.3	21.3	17.2	16.3
Northeast	7.8	14.7	8.6	17.0	13.5	12.3
South	11.4	12.3	10.8	17.5	15.5	13.5
West	12.0	12.7	12.9	16.2	10.5	12.9
Total	12.1	13.1	11.4	17.8	14.2	13.7
Number of Gang Members						
Population						
25,000 thru 49,999	34.3	37.3	42.9	42.1	42.3	39.8
50,000 thru 99,999	25.8	35.6	31.2	36.0	34.1	32.5
100,000 thru 199,999	26.5	31.7	29.4	39.2	28.2	31.0
200,000 thru high	32.7	33.6	29.9	29.2	37.0	32.5
Region						
Midwest	34.3	38.9	34.9	38.3	36.8	36.6
Northeast	26.7	38.8	37.6	36.2	27.0	33.3
South	26.8	36.3	29.3	38.0	38.6	33.8
West	26.8	27.8	32.9	33.8	30.6	30.4
Total	28.3	34.7	32.5	36.6	34.6	33.3
Number of Gang Homicides						
Population						
25,000 thru 49,999	8.2	11.9	12.6	14.9	13.1	12.1
50,000 thru 99,999	13.7	12.1	7.3	14.3	12.7	12.0
100,000 thru 199,999	14.2	18.0	15.9	20.9	17.5	17.3
200,000 thru high	35.6*	30.4*	19.7*	32.5*	30.3*	29.7
Region						
Midwest	14.2	16.1	14.4	19.9	19.0	16.7
Northeast	15.6	17.2	17.2	14.9	10.1	15.0
South	20.9	19.3	11.8	25.1	19.8	19.4
West	13.4	12.2	8.6	12.3*	12.9*	11.9
Total	16.4	16.3	12.1	18.9	16.5	16.0
N	653	802	736	702	787	

* $p < .05$; chi-square difference within year by population or region.

a given measure, by community size and region. For example, the correlation in the top left corner shows that the number of youth gangs in 2005 and the number in 2006 have a Pearson's correlation of 0.584 in communities with a population between 25,000 and 49,999. Additionally, to help with interpretation, we shaded all the correlations in Tables 3, 4, 5. The darkest shade indicates high reliability, light shade indicates moderate reliability, and no shade indicates low reliability.[6] Findings in Table 3 suggest that NYGS measures of gangs and gang members are generally very reliable across the five-year period, with little variation by population or region. The adjacent year average, over the five-year period studied, was above 0.74 for both number of youth gangs and number of gang members regardless of the population size of the jurisdiction or region the data came from.

The pattern of results for gang homicides differs from the previous two measures. The adjacent year average, over the five-year period studied, was high; however, there was substantial variation, particularly with respect to population size. Gang homicide data were the most reliable in jurisdictions with populations of 200,000 or more ($r = 0.97$), followed by jurisdictions with populations of 100,000 to 199,999 ($r = 0.79$), 50,000 to 99,999, ($r = 0.67$) and 25,000 to 49,999 ($r = 0.64$). Since gang homicides are a relatively rare event in smaller jurisdictions, we would expect a greater volatility to the annual count, which is the likely source of the lower levels of reliability for these jurisdictions. Gang homicide data were reliable regardless of region, since population size is less of a factor and is more balanced across region.

Internal Consistency Reliability Among Measures

Table 4 shows the internal consistency across the three gang measures, partitioned by population, region, and year. Generally, the data on number of gangs and gang homicides displayed low reliability. That is, the number of gangs and the number of gang member-based homicides were not associated with one another. For example, the correlation coefficients between the number of gangs and gang homicides rarely approached more than 0.5 when examined by the size of the jurisdiction. The only exception was for those jurisdictions located in the West, where three of the five-year correlation coefficients were above 0.6, and the five-year average was 0.61. The analyses did suggest that there is some reliability between the number of gangs and the number of gang members, but it varied by community

---

[6] As a general principle attaching labels or categorizing correlation coefficients is to be avoided, but it is occasionally done to increase the interpretability of findings and their readability. This is the case here. Landis and Koch (1977) noted that correlation coefficients between .81 and 1.0 are "almost perfect" and that correlation coefficients between .61 and .80 are "substantial." A review of the general literature suggests that reliability coefficients should be at least .70 (Nunnally, 1978), but are often interpreted as moderate when they are as low as .60 or somewhat lower (Taylor, 1990). As a consequence, we relied on this general framework for the interpretation of our results.

❋ Table 3

Test-Retest Reliability Among Measures by Population and Region

	2005-06	2006-07	2007-08	2008-09	Adjacent Year Average
Number of Gangs					
Population					
25,000-49,000	0.584	0.757	0.745	0.878	0.741
50,000-99,999	0.741	0.850	0.880	0.819	0.823
100,000-199,000	0.912	0.863	0.901	0.831	0.877
200,000-high	0.902	0.798	0.899	0.973	0.893
Region					
Midwest	0.944	0.762	0.841	0.849	0.849
Northeast	0.889	0.890	0.935	0.824	0.885
South	0.835	0.825	0.866	0.985	0.878
West	0.962	0.971	0.948	0.978	0.965
Number of Gang Members					
Population					
25,000-49,000	0.840	0.745	0.967	0.917	0.867
50,000-99,999	0.711	0.912	0.828	0.844	0.824
100,000-199,000	0.928	0.981	0.977	0.788	0.919
200,000-high	0.910	0.951	0.937	0.992	0.948
Region					
Midwest	0.895	0.917	0.856	0.999	0.917
Northeast	0.743	0.954	0.970	0.567	0.809
South	0.889	0.930	0.932	0.928	0.920
West	0.937	0.987	0.982	0.912	0.955
Number of Gang Homicides					
Population					
25,000-49,000	0.730	0.573	0.777	0.488	0.642
50,000-99,999	0.498	0.785	0.726	0.658	0.667
100,000-199,000	0.912	0.814	0.642	0.786	0.789
200,000-high	0.981	0.985	0.952	0.963	0.970
Region					
Midwest	0.994	0.989	0.993	0.989	0.991
Northeast	0.932	0.707	0.707	0.763	0.777
South	0.863	0.840	0.929	0.908	0.885
West	0.981	0.989	0.975	0.976	0.980

Level of Reliability	Value	Color Code
High Reliability	+ .80 thru 1.0	
Moderate Reliability	+ .60 thru .79	
Low Reliability	+ .00 thru .59	

Note. All correlations are significant at $p < .05$.

❈ Table 4

Internal Consistency Reliability Between Gang Measures by Population and Region—Pearson's R

Gangs		25,000 - 49,000	50,000 - 99,999	100,000 - 199,000	200,000 - high	Midwest	Northeast	South	West
Gang Members									
2005		0.357	0.416	0.349	0.786	0.742	0.755	0.659	0.813
2006		0.641	0.375	0.759	0.785	0.690	0.322	0.650	0.840
2007		0.463	0.206	0.527	0.821	0.570	0.535	0.670	0.792
2008		0.480	0.504	0.461	0.778	0.650	0.480	0.842	0.802
2009		0.596	0.469	0.349	0.333	0.140	0.839	0.745	0.695
5-year average		0.507	0.394	0.489	0.701	0.558	0.586	0.713	0.788
Gangs	**Member Homicides**								
2005		0.127 [a]	0.217	0.159 [a]	0.231 [a]	0.384	0.574	0.351	0.376
2006		0.152 [a]	0.287	0.208	0.688	0.735	0.678	0.554	0.696
2007		0.126 [a]	0.263	0.134 [a]	0.801	0.404	0.333	0.505	0.877
2008		0.382	0.407	0.184 [a]	0.258	0.567	0.501	0.434	0.365
2009		0.029 [a]	0.301	0.186	0.537	0.162 [a]	0.627	0.415	0.741
5-year average		0.163	0.295	0.174	0.503	0.450	0.543	0.452	0.611
Gang Members	**Member Homicides**								
2005		0.015 [a]	0.363	0.545	0.435	0.586	0.505	0.302	0.708
2006		0.255	0.412	0.295	0.931	0.263	0.515	0.767	0.808
2007		0.463	0.171	0.453	0.866	0.660	0.511	0.588	0.817
2008		0.293	0.409	0.468	0.957	0.994	0.614	0.610	0.506
2009		0.060 [a]	0.395	0.554	0.954	0.978	0.617	0.534	0.959
5-year average		0.217	0.350	0.463	0.829	0.696	0.552	0.560	0.760

Level of Reliability	Value	Color Code
High Reliability	+ .80 thru 1.0	
Moderate Reliability	+ .60 thru .79	
Low Reliability	+ .00 thru .59	

[a] = nonsignificant; all others are statistically significant at the $p < .05$.

size and geographic region. For example, looking at Table 4, with the exception of 2009, there was high reliability between the number of gangs and gang members in jurisdictions with 200,000 or more residents. The internal consistency reliability of these data in smaller jurisdictions was low. Internal consistency between the measures of the number of gangs and gang members was relatively high in the South and West ($r = 0.71$ and $r = 0.79$, respectively) across the five study years, but was relatively low in the Midwest (five-year average, $r = 0.56$) and Northeast (five-year average, $r = 0.59$). Also, based on Pearson's correlation coefficients, NYGS measures of gang members and gang homicide appear to be reliable in the largest communities (five-year average, $r = 0.83$), and, in general, in midwestern (five-year average, $r = 0.70$) and western jurisdictions (five-year average, $r = 0.76$). The correlation coefficients between gang members and gang homicide were consistently low in jurisdictions with less than 200,000 and in jurisdictions located in the Northeast and South.

Inter-Observer Reliability of Measures

As a form of inter-observer reliability, we compared the data from the NYGS to the data from the Arizona Gang Threat Assessment (AGTA)[7] for the 18 Arizona jurisdictions that participated in both surveys.[8] Table 5 displays the percentage of jurisdictions that provided a consistent response on the presence of gangs in their community and Spearman correlation coefficients related to the number of gang members reported in each survey. The use of Spearman correlations allows for a rank-order comparison of the two surveys—to what extent are the Arizona jurisdictions similarly ranked? If the rank-ordering of jurisdictions is similar across data sets, it implies the surveys are measuring the same (or similar) phenomena. We use the Spearman correlation because of a small number of observations (n = 18) and the fact that one observation point (Phoenix) would otherwise drive the analysis (see, e.g., Weisburd & Britt, 2007, pp. 400–403). By using the Spearman correlation, we are able to assess whether the two different data sources are measuring the presence of gangs in comparable ways.

The results in Table 5 suggest a moderately high level of inter-observer reliability. For the presence of youth gangs, correlation of rank order was near perfect (2008) or perfect (2007 and 2009) across the two surveys. Similarly, data regarding the number of gang members in each jurisdiction were reliable across the two surveys, with the lowest correlation in 2008 (0.786), but much higher correlations in 2007 and 2009.

---

[7] National Gang Threat Assessment data produced by the FBI for its annual report are not available to the public.

[8] In the present study the "observer" is the reporting organization. We do not know whether the same individual in the organization completed both the NYGS and the AGTA. This issue is discussed further in the discussion section.

▩ Table 5

Inter-Observer Reliability Between National Youth Gang Center Data and Arizona Gang Survey Data

	Gang Activity			Number of Youth Gang Members		
	2007	2008	2009	2007	2008	2009
Percent Reliable	88.2	100	87.5			
Spearman				0.976	0.786	0.898
n =	17	14	16	8	8	8

Level of Reliability	Value	Color Code
High Reliability	+ .80 thru 1.0	
Moderate Reliability	+ .60 thru .79	
Low Reliability	+ .00 thru .59	

Note. All values are statistically significant at $p < .05$.

▩ Discussion

The purpose of the present study was to examine the reliability of police gang data at the aggregate level through National Youth Gang Center Survey data. Specifically, we were interested in understanding whether the NYGS is a source of reliable gang data. We examined the reliability of the data in several ways. First, we assessed the extent of missing data derived from the survey. Second, we assessed the test-retest reliability of NYGS data by examining its consistency across year, and determining whether reliability varied by region and population size. We also examined the internal consistency by examining correlation coefficients between three gang measures—number of gangs, number of gang members, and number of gang homicides. Lastly, we examined the inter-observer reliability of the NYGS data by comparing NYGS data with data obtained from the Arizona Gang Threat Assessment.

We found that the NYGS measure of number of gangs had the least amount of missing/don't know responses. Furthermore, the level of missing/don't know responses for this measure was not significantly associated with population size or geographic region. Our findings were, however, somewhat mixed with respect to the NYGS's measure of number of gang members and gang homicides. For the gang member measure we found that while there was no significant difference in the number of missing responses by population size or jurisdiction, a relatively large proportion of the jurisdictions did not provide these data. For example, over the five years of data collection, about one third of jurisdictions did not provide

data on the number of gang members in their jurisdiction. This is not only problematic with respect to policymakers' access to this information, but it also has implications for the interpretation of findings. Specifically, the amount of missing data for the measure of number of gang members will necessarily result in a loss of statistical power. Researchers may be less likely to detect a relationship when one exists. Relatedly, the extent of missing data associated with this measure can bias parameter estimates. Roth (1994) notes that "missing data may bias correlation coefficients downward. The downward bias is most likely as high and/or low scores lost on either predictors or criteria restrict the variance in one variable and attenuate the correlation with another variable" (p. 538). As discussed below, the employment of missing data techniques might be necessary to address this issue.

Similarly, with respect to the NYGS's measure of gang homicides, we found that a substantial proportion of large jurisdictions and those located in the South and Midwest did not report the number of gang homicides. As such, our findings suggest that at least with respect to the gang homicide data, the data are not missing at random; population size and the region of the country are related to whether a community reports missing data or returns a "don't know" response. Researchers and policymakers should work to identify data collection strategies and the available resources that could be used to improve data reporting and reduce the amount of missing information, particularly in the areas identified above. While the quality of these data might improve over time, researchers in the meantime should examine and determine those missing data techniques that are most appropriate for these data. While list- or pairwise deletion might be the most commonly employed missing data techniques, they would not be well suited for some of the items included in the NYGS because of the relatively high proportion of missing data and because the data associated with certain items (i.e., gang homicides) were found to be strongly correlated with such factors as region and population size—suggesting that list- or pairwise deletion might result in a biased sample. Instead, multiple imputation-based procedures should be further examined for their utility in addressing the problem of missing data for the measures of number of gang members and number of gang homicides. This is because multiple imputation procedures take into consideration "information about the data both observed and unobserved" and replace missing values with plausible values (Meng, 1994, p. 538; see also Rubin, 1987).

Further, the analyses showed that NYGS data on the number of gangs, gang members, and gang homicides obtained from law enforcement agencies across the United States, regardless of the size of the community (except for gang homicides in jurisdictions serving fewer than 200,000 people) or region in which the community is located, are fairly robust and are generally reliable enough to be used by policymakers and academics alike. While these data have been available for well over 10 years, their use by policymakers and academics has been relatively limited. These data provide policymakers with an opportunity to make data-driven decisions about resource allocation and national-level decisions regarding the

implementation of gang prevention, intervention, and suppression programming. The data also provide academics with the opportunity to examine problems related to gangs and gang membership at the macro-level. In the end, our findings suggest that these data permit researchers to examine the scope of gang problems, and the factors that influence or impact gang problems, across the United States.

With the above said, our findings suggest that our three primary measures—gangs, gang members, and gang homicides—might not measure the same underlying phenomena. Specifically, the number of gangs in a community was not strongly related to the number of gang members or gang homicides, and the number of gang members in a community was not related to the number of gangs or gang homicides. This finding does not necessarily suggest that these data are unreliable; instead, this finding suggests that these three measures might be tapping into different gang phenomena at the community level. For example, the relationship between the number of gang members and the number of gangs in a community might reflect the nature of the gang problem in the specific community. Prior research notes that gangs vary in structure (Klein & Maxson, 2006), and gang structure can determine the size of the gang and thus the correlation between the number of gangs and the number of members. Stated more directly, it might be community-level or gang-level variation that accounts for low correlation between number of gangs and members. Similarly, the number of gang homicides in a community is, in part, the product of gang processes. The number of rivalries and the mechanisms for dealing with rivalries as well as the internal gang processes can contribute to the number of gang homicides, irrespective of the number of gangs and gang members (Klein & Maxson, 2006). The current analysis suggests that the factors that produce a large number of gang members in a given community might not be the same factors that are associated with more gang homicides or the presence of multiple gangs. Researchers should build on this finding to determine how the causal mechanisms, gang structures, and gang processes are related to each of these three outcomes.

Our findings are limited in at least four ways. The first is that the current study did not control for the agencies' capacity to collect, maintain, or disseminate gang data. Prior research has found that police agencies have varying capacity to engage in these efforts. Police agencies that have an established gang unit or repeat offender unit, or who participate in a gang task force most likely have increased capacity to collect, maintain, and disseminate such data. Accordingly, these agencies might be more likely to provide more reliable data to the National Youth Gang Center. Decker and Pyrooz (2010) found such a result in their examination of the reliability and validity of gang homicide data obtained from large police organizations. While the focus of this paper was not on the internal structure of police organizations and its impact on the reliability of data, research in the future should examine the impact of organizational capacity to collect such gang data on the reliability and validity of that data.

Second, and related to the above, we did not examine whether those who filled out the NYGS or the AGTA had the knowledge to answer the questions accurately.

In the future it would be important to know the means in which the respondents acquired the data they provided on behalf of the organization to understand the extent to which data were systematically collected or were simply the result of the respondents' "best guess." Also, prior research has determined that data collected from a single individual within an organization can result in measurement error (Weiss, 1997). Future research is needed to better understand how police organizations collect, maintain, and disseminate the information they provide for such organizational surveys.

Third, with respect to our examination of inter-observer reliability, our sample of Arizona agencies was small and our findings could be a reflection of the data collection practices in Arizona. While comparing the AGTA data to NYGS data provides some context for understanding the inter-observer reliability of NYGS data, the findings should not be generalized to other jurisdictions outside of Arizona. Future research should examine the inter-observer reliability of NYGS data by comparing them to other data collection initiatives in other states across the nation. Last, definitional issues inevitably impacted the survey results. Barrows and Huff (2009) show that there are many variations in how states determine who is a gang member. Agencies might have relied on their internal reporting systems to answer the NYGS, and these systems are susceptible to changes in the administrative or legal definition of a gang member in their jurisdiction.

In summary, our analyses of five years of National Youth Gang Survey (NYGS) data indicate that these data are fairly reliable. While reliability is necessary for establishing validity, it is not sufficient. In the future researchers should further examine the validity of these data. NYGS data has been used relatively rarely (for exceptions see Pyrooz et al., 2010; Pyrooz, 2011; Wells & Weisheit, 2001), and we are hopeful about the many potential uses for these data to examine a variety of issues surrounding community-level variation in gang-related phenomena. Additionally, we are optimistic about the potential of these data to be used to examine the impact of broader national-level policies aimed at reducing the number of gangs and gang members in communities across the United States. Our optimism extends to the value that state-based law enforcement gang surveys can have in informing state- and local-level policies aimed at addressing local gang problems when they use a methodology similar to the NYGS. We are, however, somewhat less confident about the potential use of gang homicide data at this time. Our primary concern lies with the fact that it appears that gang homicide data are missing systematically by size of the agency and are significantly more likely to be unreliable when obtained from an agency serving a population of less than 200,000 persons. Further research on these issues should be performed as these data have the potential to play a major role in our understanding of gang homicide.

❀ References

Allison, P. D. (2001). *Missing data*. Thousand Oaks: Sage.

Barrows, J., & Huff, C. R. (2009). Gangs and public policy: constructing and deconstructing gang databases. *Criminology & Public Policy, 8*(4), 675–704.

Burrell, S. (1990). Gang evidence: Issues for criminal defense. *Santa Clara Law Review, 30*, 739–790.

Chesney-Lind, M., Rockwell, A., Marker, N., & Reyes, H. (1994). Gangs and delinquency. *Crime, Law, and Social Change*, *21,* 201–228.

Crew, J. (1997). *Memo to Orange County District Attorney's Office*. San Francisco: American Civil Liberties Union.

Curry, G. D., Ball, R., Fox, R., & Stone, D. (1992). *National assessment of law enforcement anti-gang information resources*. Washington, DC: National Institute of Justice.

Curry, G. D., Ball, R. A., & Fox, R. J. (1994). *Gang crime and law enforcement recordkeeping*. Washington, DC: National Institute of Justice.

Decker, S., & Pyrooz, D. (2010). On the validity and reliability of gang statistics. *Homicide Studies, 14*(4), 359–376.

Egley, A., Howell, J. C., & Moore, J. P. (2010). *Highlights of the 2008 National Youth Gang Survey*. Washington, DC: Department of Justice, Office of Justice Programs, Office of Juvenile Justice and Delinquency Prevention.

Fox, J. (2004). Missing data problems in the SHR imputing offender and relationship characteristics. *Homicide Studies*, *8,* 214–254.

Hagedorn, J. H. (1990). Back in the field again: Gang research in the nineties. In C. R. Huff (Ed.), *Gangs in America* (pp. 240–262). Newbury Park: Sage.

Hassell, K., & Zhao, J. (2003). Structural arrangements in large municipal police organizations. *Policing, 26*(2), 231–250.

Horton, N., & Lipsitz, S. (2001). Multiple imputation in practice. *American Statistician, 55*, 244–254.

Katz, C. M. (2003). Issues in the production and dissemination of gang statistics: An ethnographic study of a large Midwestern police gang unit. *Crime and Delinquency, 49*(3), 485–516.

Katz, C. M., Maguire, E., & Roncek, D. (2002). The creation of specialized police gang units: Testing contingency, social threat, and resource-dependency explanations. *Policing: An International Journal of Police Strategies and Management, 25*(3), 472–506.

Katz, C. M., & Schnebly, S. (2011). Neighborhood variation in gang member concentrations. *Crime & Delinquency*, *57*(3), 377–407.

Katz, C. M., & Webb, V. J. (2006). *Policing gangs in America*. Cambridge University Press: New York.

Katz, C. M., Webb, V. J., Gartin, P. R., & Marshall, C. E. (1997). The validity of self-reported marijuana and cocaine use. *Journal of Criminal Justice, 25*(1), 31–42.

Katz, C. M., Webb, V., & Schaefer, D. (2000). The validity of police gang intelligence lists: Examining differences in delinquency between documented gang members and non-documented delinquent youth. *Police Quarterly*, *3*(4), 413–437.

Kennedy, D. M. (2009). Comment on "Gangs and public policy: constructing and deconstructing gang databases." *Criminology & Public Policy, 8*(4), 711–716.

King, W. (2000). Measuring police innovation: Issues and measurement. *Policing, 23*(3), 303–317.

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

Klein, M. W. (2009). Street gang databases: A view from the gang capitol of the United States. *Criminology & Public Policy, 8*(4), 717–722.

Klein, M. W., & Maxson, C. L. (2006). *Street gang patterns and policies*. Oxford: Oxford University Press.

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics, 33*(1), 159–174.

Langton, L. (2010). *Gang units in large local law enforcement agencies, 2007*. Washington DC: U.S. Department of Justice, Office of Justice Programs, Bureau of Justice Statistics.

Langworthy, R., & Travis, L. (1994). *Policing in America*. New York: MacMillian.

Maguire, E., & Katz, C. M. (2002). Loose coupling, sensemaking, and community policing in American police agencies. *Justice Quarterly, 19*(3), 501–534.

Maxfield, M. G., & Babbie, E. R. (2011). *Research methods for criminal justice and criminology* (6th Ed.). Belmont, CA: Wadsworth Cengage Learning.

Maxson, C. L. (1999). Gang homicide: A review and extension of the literature. In M. D. Smith & M. A. Zahn (Eds.), *Homicide: A sourcebook of social research* (pp. 239–254). Thousand Oaks, CA: Sage.

Maxson, C. L., & Klein, M. W. (1990). Street gang violence: Twice as great, or half as great? In C. R. Huff (Ed.), *Gangs in America* (pp. 71–100). Newbury Park, CA: Sage.

Maxson, C. L., & Klein, M. W. (1996). Defining gang homicide: An updated look at member and motive approaches. In C. R. Huff (Ed.), *Gangs in America* (2nd Ed., pp. 2–20). Newbury Park, CA: Sage.

McCorkle, R., & Miethe, T. (1998). The political and organizational response to gangs: An examination of a "moral panic" in Nevada. *Justice Quarterly, 15*(1), 41–64.

Meehan, A. (2000). The organizational career of gang statistics: The politics of policing gangs. *The Sociology Quarterly 41*(3), 337–370.

Meng, X. L. (1994). Multiple imputation inferences with uncongenial source of input. *Statistical Science, 9*(4), 538–573.

Miller, W. B. (1982). *Crime by youth gangs and groups in the United States*. Washington DC: U.S. Department of Justice, Office of Juvenile Justice and Delinquency Prevention.

National Youth Gang Center. (2008). *National Youth Gang Survey Analysis*. Available online at http://www.nationalgangcenter.gov/Survey-Analysis

National Youth Gang Center. (2009). *National Youth Gang Survey Analysis*. Available online at http://www.nationalgangcenter.gov/Survey-Analysis

Nunnally, C. (1978). Psychometric theory. New York: McGraw-Hill.

Pyrooz, D. C. (2011). Structural covariates of gang homicide in large U.S. cities. *Journal of Research in Crime and Delinquency*. Prepublished at Online First, DOI:10.1177/0022427811415535.

Pyrooz, D. C., Fox, A. M., & Decker, S. H. (2010). Racial and ethnic heterogeneity, economic disadvantage, and gangs: A macro-level study of gang membership in urban America. *Justice Quarterly, 27*(6), p. 867–892.

Roth, P. (1994). Missing data. *Personnel Psychology, 47*(3), 537–560.

Rubin, D. B. (1987). Multiple imputation for nonresponse in surveys. New York: J. Wiley & Sons.

Short, J. F. (2009). Gangs, law enforcement, and the academy. *Criminology & Public Policy, 8*(4), 723–732.

Steel, B., & Lovrich, N. (1987). Equality and efficiency tradeoffs in affirmative action—Real or imagined? The case of women in policing. *The Social Science Journal, 24*(1), 53–70.

Taylor, R. (1990). Interpretation of the correlation coefficient: A basic review. *Journal of Diagnostic Medical Sonography*, 6, 35–39.

Trochim, W. M. (2006). *The research methods knowledge base* (2nd Ed.). Available online at http://www.socialresearchmethods.net/kb/

Walker, S., & Katz, C. M. (2005). Less than meets the eye: Police department bias crime units. *American Journal of Police, 14* (1), 29–48.

Webb, V. J., & Katz, C. M. (2003). Policing gangs in an era of community policing. In S. Decker (Ed.), *Policing gangs and youth violence* (pp. 17–49). Wadsworth: Belmont.

Webb, V. J., & Katz, C. M. (2006). A study of police gang units in six cities. In A. Egley, C. Maxson, J. Miller, & M. Klein (Eds.), *The modern gang reader* (3rd Ed., pp. 349–360.) Los Angeles: Roxbury Publishing.

Webb, V., Katz, C. M., & Decker, S. (2006). Assessing the validity of self-reports by gang members: Results from the Arrestee Drug-Abuse Monitoring program. *Crime & Delinquency, 52*(2), 232–252.

Weisburd, D., & Britt, C. (2003). *Statistics in criminal justice* (2nd Ed.). Belmont: Wadsworth.

Weisburd, D., & Britt, C. (2007). Statistics in criminal justice (3rd Ed.). Belmont, Wadsworth.

Weiss, A. (1997). The communication of innovation in American policing. *Policing, 20,* 292–310.

Wells, L. E., & Weisheit, R. A. (2001). Gang problems in nonmetropolitan areas: A longitudinal assessment. *Justice Quarterly*, *18*(4), 801–833.

Wilson, J. Q. (1968). *Varieties of police behavior*. Cambridge: Harvard University Press.

Zatz, M. (1985). Los cholos: Legal processing of Chicano gang members. *Social Problems, 33*(1), 13–30.

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

Charles M. Katz is Watts Family Director of the Center for Violence Prevention and Community Safety and is a Professor in the School of Criminology and Criminal Justice at Arizona State University. He received his Ph.D. from the University of Nebraska at Omaha in 1997. His research interests include gangs and gang control, policing, and quality of gang data and police organizational theory.

Andrew M. Fox is an Assistant Professor of Criminal Justice and Criminology at the University of Missouri - Kansas City. He is expected to receive his Ph.D. from Arizona State University in 2012. His research interests include social network analysis, gangs, drug markets, and communities.

Chester L. Britt is Dean of the School of Criminology and Criminal Justice at Northeastern University. His research interests include quantitative methods, criminological theory, and decisionmaking in the criminal justice system.

Phillip Stevenson is the Director of the Arizona Statistical Analysis Center, which is the research arm of the Arizona Criminal Justice Commission. The Arizona Statistical Analysis Center conducts research of practical value to criminal justice practitioners and policymakers. Dr. Stevenson serves as Co-Chair of Arizona's National Instant Criminal Background Check System Task Force and is Vice Chair of the Global Justice Information Sharing Initiative Privacy and Information Quality Working Group. He is also the current President of the Justice Research and Statistics Association. He received his bachelor's and master's degrees in sociology from the University of Arizona and his Ph.D. in sociology from Loyola University Chicago.