

Estimating the Impact of Classification Error on the “Statistical Accuracy” of Uniform Crime Reports

James J. Nolan · Stephen M. Haas · Jessica S. Napier

© Springer Science+Business Media, LLC 2011

Abstract This paper offers a methodological approach for estimating classification error in police records then determining the statistical accuracy of official crime statistics reported to the Uniform Crime Reporting (UCR) program. Classification error refers to the mistakes in UCR statistics caused by the misclassification of criminal offenses, for example recording a crime as aggravated assault when it should have been simple assault. Statistical accuracy refers to the estimated true total of each crime type based on cancelling effect of undercounting and overcounting crime due to misclassifications. The population for the study consists of the 12 largest municipal police agencies in a mostly rural southeastern state. Based on a sample of 2,663 records, the authors illustrate the impact of classification error on the total population of reported offenses. Misclassifications result in overcounting and undercounting certain crimes. The true number of each crime type, as well as the aggregate Index Crime, Violent Crime, and Property Crime totals, is estimated based the evaluation of offsetting misclassifications. The findings show that certain UCR crime categories are greatly undercounted while others are overcounted. The index crime and violent crime totals are also significantly undercounted; however, when simple assault is added to the index and violent crime categories, the error in these aggregate numbers is reduced to less than 1%. The results provide a benchmark for assessing the statistical accuracy of the UCR data.

Keywords Classification error · Statistical accuracy · Uniform Crime Reports (UCR) · Official crime statistics · National incident-based reporting system (NIBRS)

J. J. Nolan (✉)
Department of Sociology and Anthropology, West Virginia University,
P.O. Box 6326, 309 Knapp Hall, Morgantown, WV 26506-6326, USA
e-mail: jim.nolan@mail.wvu.edu

S. M. Haas · J. S. Napier
West Virginia Division of Justice and Community Services, Charleston, WV, USA

Introduction

The purpose of this article is to explore the phenomenon of classification error in the Uniform Crime Reports (UCR). Classification error occurs when criminal offenses are misclassified by police before they are entered into the UCR database. This misclassification results in over counting some offenses while undercounting others. Classification error in UCR, whether intentional or accidental, results in the inflation or deflation of crime in UCR statistics.

The UCR Program is a national initiative involving more than 17,000 city, county, state and federal law enforcement agencies which voluntarily report crime data to the Federal Bureau of Investigation (FBI 2007). The main objective of the UCR Program is to generate a valid set of crime data for use in law enforcement administration, operation, and management. Since the inception of the UCR Program more than 70 years ago, police records have been the primary source of data for the national crime statistics. Although the quality of police records has often come into question, UCR has been found to be a valid indicator of the “index” crimes (Gove et al. 1985). Undeniably, the information gathered and reported through the UCR Program has become a social indicator for the nation and is routinely used by “criminologists, sociologists, legislators, municipal planners, the media, and other students of criminal justice for varied research and planning purposes” (FBI 2007, para. 1).

Such widespread use of data has underscored the importance of ensuring its accuracy (Nolan et al. 2006). The Federal Bureau of Investigation has acknowledged that the integrity and accuracy of the UCR depends upon the data entered by the participating agencies (Akiyama and Propheter 2005); thereby, suggesting that attention should be given to the errors produced by law enforcement officers and the effects these errors have on the quality of the UCR data. While considerable attention has been focused on errors associated with victim reporting, police recording, and the handling of missing data (Lynch and Jarvis 2008; Hart and Rennison 2003; Maltz 1999), few, if any, published studies have been undertaken to assess crime classification error.

Once classification errors have been identified and measured in each crime type, the statistical accuracy of the crime data can be determined. By *statistical accuracy* we refer to the numerical difference (or lack thereof) between the *recorded number* of crimes and *true number*¹ of crimes reported to UCR. The true number of crimes in each criminal offense category is an estimate that takes into account the offsetting effect of undercounting and over counting because of misclassification.

Since there is no uniform method for identifying and assessing the impact of classification error in the US (Lynch and Jarvis 2008; Maltz and Targonski 2002; Nolan et al. 2006; Woodard 1992), the central purpose of this study is to introduce a methodology for assessing the “statistical accuracy” of crime estimates produced by the UCR Program. In an effort to illustrate the merit of this methodology, this study examines a random sample of crimes reported to the UCR Program from 12 cities and towns in a southeastern state. An analysis is conducted to identify the under- and over-counting of offenses across various crime types. Error rates are applied to calculate point estimates with confidence intervals of the *true* number of crimes in each offense category. Again, the true number is not the

¹ By “true number” we are only referring to the crimes that have made it into the police database—i.e., the source of the UCR. This in no way means the true number of criminal offenses that actually occurred in the police jurisdiction. As we explain later in this paper, many crimes go unreported by victims or are not recorded by police when they are reported.

number of crimes that had actually occurred in the jurisdiction, but the true number of each offense type that made its way into the police database. Both the sources of classification error and the most common reasons for misclassifications are examined. The following section offers a brief overview about classification error and its relationship to the accuracy of official crime statistics.

Literature Review

It is well known in the fields of criminology and sociology that the UCR is a “statistical program,” meaning that it is *not* an actual accounting of all crimes that occur in a jurisdiction. Moreover, UCR is *not* a sample either. Instead, UCR provides a full accounting of crimes that are recorded by police. Since its inception in the late 1920s, the UCR Program has been a valuable resource to police and criminologists alike in their efforts to understand the nature and extent of crime locally and nationally. Although UCR is not produced by sampling, the program is considered statistical because adjustments are made for missing or erroneous data.² Since it is a statistical program, the value of UCR is not contingent on the elimination of all errors (Nolan et al. 2006). Instead, it is the “error structure” of UCR that is important, please refer to Fig. 1.

Block A in Fig. 1 represents all crimes that occur within a particular jurisdiction. This includes all behavior that fits the definition of a crime whether it is reported or not. Moving all the way to the right of Fig. 1 is block D “police records.” Block D represents all crimes that occurred in a jurisdiction which were recognized as crimes and reported by victims and recorded by police. Once the police create a report of an incident the crime makes its way into the police database, i.e., block D. Of course D will be less than A because sometimes the victims do not recognize that a crime occurred (A'), and many times and for a variety of reasons victims decide not to report crimes to the police (B'), and quite often the police do not write reports of incidents (C') even when they are reported to them. Given all of the places in the reporting process where crimes that occurred in A can escape being counted in D, D has still been found to be a valid measure of A—with some caveats (Gove et al. 1985).

In addition to errors/omissions at A' , B' and C' in Fig. 1, error is introduced into UCR totals by the use of estimation and imputation procedures when data are prepared for publication. We represent these errors as D_1 . It is not uncommon, to have within a single jurisdiction missing months of crime reports because of backlogs or corrupted files. Or within a state where multiple jurisdictions contribute to the state total, missing agency data will affect the aggregate totals and will impact state and national trend analyses among other things (Lynch and Jarvis 2008; Maltz 1999). In Fig. 1, D_2 represents crime classification error which is the focus of this study and will be explained in much detail below.

To date, most studies of the UCR error structure have focused on victim-reporting (B') and police-recording (C') errors. The literature on crime reporting is replete with examples of these types of errors (see Bachman 1993; Black 1974; del Frate and Goryainov 1994; Gove et al. 1985; Greenberg and Ruback 1992; Schwind and Zwenger 1992; Shah and Pease 1992; Skogan 1976; Warner and Pierce 1993; Wexler and Marx 1986; White and Mosher 1986 as examples). For example, in a study of the 2000 National Crime

² We refer to UCR as a “statistical” program not to indicate that crime totals are obtained by sampling. Instead we use this terminology to make the distinction between the use of statistical methods and “accounting” or “book keeping” methods. The UCR does not just compile the data that get reported. It uses statistical techniques to address missing data and for overall quality control.

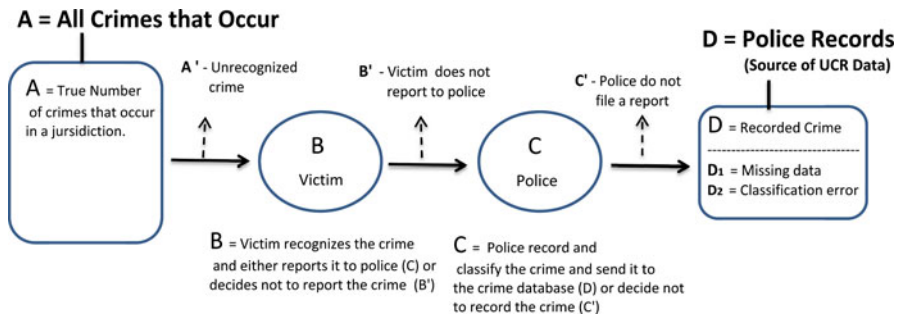


Fig. 1 Error structure in UCR data

Victimization Survey (NCVS), Hart and Rennison (2003) found that only 4 in 10 property crimes and 6 in 10 violent crimes were never reported to police.

To our knowledge there have been no published studies of classification error (D_2) in the UCR statistics, so little is known about the degree to which this type of error influences the statistical accuracy of aggregate crime statistics. We also do not know how to deal with classification error in UCR once it is discovered (Nolan et al. 2006). For these reasons, the current study was initiated.

Sources of Classification Error

A classification error occurs when police officers record the facts of an incident correctly, but misclassify the crime type (Nolan et al. 2006). Theoretically, classification errors could occur for multiple reasons, including simple mistakes by officers, bad report writing habits passed down from officer to officer, the deliberate downgrading of particular crimes to reduce the crime rate (Forero 2000; McCoy et al. 1998; Ombudmans Victoria 2009; Rashbaum 2003), and from automation problems (Dodge 1990).

Regardless of the reason for classification error, clearly it can have a substantial impact on the accuracy of the published UCR statistics and on the decisions made and conclusions drawn by consumers of the data. In the following section we describe two basic ways of assessing classification error, i.e., through “record” accuracy and “statistical” accuracy.

Record Versus Statistical Accuracy

Classification error can be assessed in two ways, through (1) record accuracy and (2) statistical accuracy (Nolan et al. 2006). Record accuracy provides only a partial look at the impact of classification error on a particular crime type while statistical accuracy provides a comprehensive assessment of all of the classification errors in the police records. Record accuracy refers to the estimate of classification error in specific crime types viewed alone, such as when police want to know the accuracy of their larceny statistics. Record accuracy (R_i) is calculated by Eq. 1:

$$R_i = \frac{o_i}{n_i} N_i, \tag{1}$$

where o_i , the number of classification errors in a particular crime category; n_i , the sample of reported crimes in a particular category of crime; N_i , the total number of reported crimes in a particular category of crime.

To estimate the record accuracy of UCR statistics, one could draw a random sample of police reports in a particular category of crime, for example assault or robbery. Once the sample was drawn the researcher could read the narratives and decide whether the classifications of the crimes were correct. The total number of errors divided by the number of records in the sample provides the error rate, see Eq. 1. The error rate, then, can be multiplied by the total number of records in population (N_i) to obtain a point estimate of the number of record errors that are in that specific crime type in the police database. As one can see, this type of examination provides an assessment of only the over counted crime in that crime category.

Statistical accuracy, on the other hand, deals with both over counting and undercounting errors. Statistical accuracy refers to the accuracy of the crime totals after all crime types have been examined and offsetting misclassifications have been considered. Since some of the misclassifications result in over counting UCR crimes while others result in undercounting them, the correct UCR number can be obtained only by considering the canceling effect of the two types of errors—over counting and undercounting. Please refer to Fig. 2.

Figure 2 is a depiction of the process of calculating statistical accuracy in UCR. There are 5 categories of records that must be considered in order to determine statistical accuracy:

- A The UCR crime category of interest.
- B The UCR crime categories that were erroneously reported as A, resulting in an over count of crime A.
- C The UCR crime categories (not A) that should have been recorded as A, resulting in an undercount of crime A.
- D The non UCR incidents that should have been reported as A, resulting in an undercount of crime A.
- E The non UCR incidents that were erroneously reported as A, resulting in an overcount of crime A.

An example may help clarify the difference between record accuracy and statistical accuracy (see Table 1). Let us suppose the police want to assess the accuracy of their UCR records and, in particular, they are interested in the crime category “simple assault.” In this hypothetical situation, a stratified random sample of 646 records was pulled from the police database. The total number of records for the period studied (e.g., 1 year) was 41,320.

Fig. 2 Illustration of record and statistical accuracy

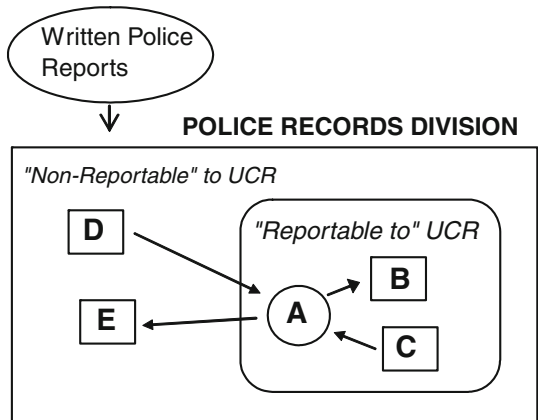


Table 1 Hypothetical study of police reports for record and statistical error

Category	Sample size (n_i)	Total records (N_i)
Murder	$n_m = 1$	$N_m = 1$
Rape	$n_r = 15$	$N_r = 20$
Robbery	$n_{rob} = 30$	$N_{rob} = 300$
Aggravated assault	$n_{aa} = 50$	$N_{aa} = 500$
Burglary	$n_b = 100$	$N_b = 3,000$
Theft	$n_t = 100$	$N_t = 5,000$
M. V. theft	$n_{mv} = 100$	$N_{mv} = 1,000$
Arson	$n_a = 50$	$N_a = 499$
Simple assault	$n_{sa} = 100$	$N_{sa} = 1,000$
Reportable crimes (sum of above)	$n_{rep} = 546$	$N_{rep} = 11,320$
Non-reportable crimes	$n_{nr} = 100$	$N_{nr} = 30,000$
Total sample	$n_{total} = 646$	$N_{total} = 41,320$

Of these records, 30,000 are non-UCR incidents and 11,320 are crimes.³ The database is partitioned into nine UCR categories (simple assault is included) and one category for non UCR incidents. A simple random sample is then drawn from each of the 10 categories. This sample size totals 646 records. Table 1 provides the necessary data.

The following notations (A, B, C, D, E) apply to Fig. 2.

- A 100 reported simple assaults selected randomly from universe of simple assaults.
- B 9 crimes reported as “simple assault” but were classified as another *reportable* (UCR) crime. In this case all 9 were reported as aggravated assault.
- C 11 simple assaults that were actually classified as some other *reportable* (UCR) crime. In this example, 5 had been classified as aggravated assault and 6 had been classified as robbery.
- D 1 crime categorized as *non-reportable* (non UCR) but which was really simple assault.
- E 10 crimes recorded as “simple assault” that were actually *non-reportable* (non UCR).

Calculating Record Error

As described in Eq. 1, record accuracy can be measured by dividing the overcount of crime in a particular category by the number reported. In the situation presented above, category B and category E represent overcounts of simple assault. Therefore, the record error in this sample is calculated as follows according to Eq. 1:

$$\text{Record error } R_i = \left(\frac{B + E}{A} \right) 100 = \left(\frac{19}{100} \right) 100 = 19\%$$

³ Not all crimes are UCR crimes and not all police reports deal with specific crimes. “Reportable UCR crimes” include murder, rape, robbery, assault (simple and aggravated), burglary, motor vehicle theft, larceny/theft, and arson. Non UCR incidents include reports on other criminal offenses not included in UCR or police reports that document situations such as “suspicious persons” or “ongoing neighborhood disputes” that are not crimes.

Calculating Statistical Error

Statistical accuracy is a measuring of both the over counting and under counting of a particular crime category. Statistical accuracy is measured in two steps. The first step involves the calculation of the true number of crimes in a particular crime category. This actual crime total is denoted by T_i . In this case T_{sa} is calculated by subtracting over counts from the reported number of simple assaults and then adding the undercounts. Applying the notations from Table 1 and Fig. 2 in this example, T_{sa} can be calculated according to Eq. 2:

$$\begin{aligned}
 T_{sa} &= N_{sa} - \left(\left(\frac{B}{n_{sa}} N_{sa} \right) + \left(\frac{E}{n_{sa}} N_{sa} \right) \right) + \left(\left(\frac{C_{aa}}{n_{aa}} N_{aa} \right) + \left(\frac{C_{rob}}{n_{rob}} N_{rob} \right) + \left(\frac{D}{n_{nr}} N_{nr} \right) \right) \\
 &= 1000 - \left(\left(\frac{9}{100} \right) 1000 + \left(\frac{10}{100} \right) 1000 \right) + \left(\left(\frac{5}{50} \right) 500 + \left(\frac{6}{30} \right) 300 + \left(\frac{1}{100} \right) 30000 \right) \\
 &= 1000 - (90 + 100) + (50 + 300) = 1160
 \end{aligned}
 \tag{2}$$

B and E represent *overcounts* of category A (denoted above as N_{sa}); B are overcounts that come from UCR crimes and E are overcounts that come from non UCR crimes. C and D are *undercounts* of crime category A. In category C are crimes that were reported as some other UCR crime, but should have been reported as category A. In category D are non UCR crimes that should have been recorded as category A (simple assault in this example).

In the case of statistical accuracy, it is important to note that the errors found in the sample in each category (B through E) are actually error proportions. B and E are proportions of the simple assault sample that actually belong in other categories. C is the proportion of other UCR crimes that are actually simple assault. D is the proportion of non-reportable crimes that were actually simple assaults.

The second step in determining statistical accuracy is to divide the reported number of crimes (N_i) by the estimated true number of crimes in the crime category (T_i). Subtracting 1 and multiplying by 100 produces the percent of under or over count. The statistical error rate in this example is calculated as follows in Eq. 3. The result in this hypothetical case is a 14% undercount of simple assault.

$$\frac{N_{sa}}{T_{sa}} - 1 = \frac{1000}{1160} - 1 = 0.86 - 1 = -0.14
 \tag{3}$$

Methodology

The methods used to assess classification error and calculate statistical accuracy in this study involved the identification of specific crime categories for examination and three distinct stages: (1) pre-sampling, (2) selecting and reviewing sampled records, and (3) calculating statistical accuracy. Each of these stages is described below.

Pre-Sampling

The pre-sampling stage involved two steps: (1) the partitioning of records and (2) the calculation of the appropriate sample size. The 12 largest municipal police departments in

the state composed the population for this study. The combined population in these cities is 285,898 and the total number of incidents reported to and recorded by police is 31,003. This number does not include the number of reports that render the incidents as “unfounded.” When an “unfounded” report is written it is often after an initial investigation reveals that the crime really did not occur. The number of incidents (31,003) also does not include “general incident” reports which are written when police take some action and want to document it, but it is not related to a specific crime. When we add “unfounded” and “general incident” reports the population of reports grows to 31,084.⁴ Table 2 describes the resident populations and number of crimes reported within each of these 12 agencies in 2002 (i.e., the largest municipal police departments in a largely rural southeastern state⁵). From the group of 12 police departments, three were randomly selected to participate. The records in each of these three agencies were partitioned into 15 categories.

Table 3 describes these three police departments and the 15 categories, including unfounded and general incident reports and the estimated sample size for each category and participating police agency. The column “*N*” under each agency describes the total number of crime reports within each of these three participating agencies, while the “*n*” column provides the sample sizes by crime category for each of the three agencies. The “*N*” under the “Stratum Total” column represents the number of reports in each crime category, for all 12 municipal police agencies, representing the total population for this study (31,084 incidents). Similarly, the “*n*” under the “Stratum Total” column is the total number sampled from the three participating agencies (2,689 incidents).

In order to establish point estimates of statistical accuracy, a sample of records from each of the 15 reporting categories was drawn. Prior to drawing the sample, however, it was necessary to decide on a desired level of confidence and an acceptable rate of error. The sample size for each stratum (category) was calculated according to Eq. 4.

$$n = \frac{k^2 NPQ}{k^2 PQ + NE^2}, \quad (4)$$

where *n*, desired sample size; *k*, 95% confidence level reflected in the *z* score 1.96; *N*, number in the population from which the sample is to be drawn; *P*, the estimated error rate in the population; *Q* = 1 – *P*; *E*, the range of error in the estimate such as ±0.03 (3%).

The estimated sample size for each reporting category and participating police agency is provided in Table 3. Additionally, Table 4 identifies the desired sample size for each crime category based on the estimate of error expected to be found.

Sample Selection and Review of Records

An automated random sample generator was used to select the records for this study. With the assistance of police agency representatives, hard copies of all reports in the sample

⁴ In the state where this research was conducted, “unfounded” reports and “general incident reports” are not tallied and so the population of these reports is unknown. Officers tell us that these are reports are rarely written. We were able to find 31 general incident reports and 50 unfounded reports and we are, therefore, treating these reports as the population rather than as a sample (see Table 2).

⁵ It is not required by law for states to submit data to the UCR. However, in the state where this study occurred, it is required for all law enforcement agencies to submit NIBRS data to the Bureau of Uniform Crime Reports at the state police. This state received its NIBRS certification from the FBI in 1999 and its estimated coverage is 100% of the population.

Table 2 Incidents reported for the twelve largest municipal police agencies by offense category

	Agency 1	Agency 2	Agency 3	Agency 4	Agency 5	Agency 6	Agency 7	Agency 8	Agency 9	Agency 10	Agency 11	Agency 12	Total
Resident population	14,918	51,291	20,338	15,633	53,230	13,342	19,028	11,410	25,713	17,193	32,980	10,822	285,898
Total incidents	2,915	6,539	1,279	1,290	8,984	95	1,584	1,360	2,552	447	3,338	620	31,003 ^b
Arson	6	30	3	5	47	0	4	6	8	2	25	1	137
Aggravated assault	55	106	17	33	306	0	29	26	73	4	17	7	673
Simple assault ^a	339	913	262	386	1,306	12	22	210	228	18	494	90	4,480
Burglary	173	714	82	64	840	9	103	82	222	7	194	32	2,522
Murder	0	4	0	0	7	0	1	2	1	0	0	0	15
Other homicide	0	0	0	0	0	0	0	0	0	0	0	0	0
Larceny	822	1,942	280	181	2,650	27	381	247	718	96	817	267	8,428
M. V. thefts	76	311	13	18	461	5	46	21	39	5	49	12	1,056
Robbery	31	91	7	5	163	0	20	12	20	0	16	1	366
Rape	2	43	4	6	27	0	3	4	13	0	22	1	125
Other sex offenses	17	14	2	10	39	0	7	6	9	0	34	4	142
Other Group A	426	1,581	278	162	1,786	32	365	277	886	32	688	138	6,651
Other Group B	968	790	331	420	1,352	10	403	467	335	283	982	67	6,408

^a Simple assault includes intimidation

^b The total number of incidents (31,003) does not include unfounded or general incident reports

Table 3 Distribution of population and sample incidents by agency and offense category

	Agency 5		Agency 7		Agency 9		Stratum ^c	Total
	<i>N</i>	<i>n</i>	<i>N</i>	<i>n</i>	<i>N</i>	<i>n</i>	<i>N</i>	<i>n</i>
Arson	47	47	4	4	8	8	137	59
Aggravated assault	306	166	29	26	73	18	673	210
Simple assault ^a	1,306	118	222	60	228	127	4,480	305
Burglary	840	226	103	43	222	175	2,522	444
Murder	7	7	1	1	1	1	15	9
Other homicide	0	0	0	0	0	0	0	0
Larceny	2,650	375	381	61	718	242	8,428	678
M. V. thefts	461	94	46	38	39	32	1,056	164
Robbery	163	100	20	18	20	17	366	135
Rape	27	16	3	3	13	13	125	32
Other sex offenses	39	36	7	7	9	7	142	50
Other Group A	1,786	168	365	69	886	62	6,651	299
Unfounded	31	31	— ^b	0	—	0	31	31
General incidents	48	48	2	2	—	0	50	50
Group B	1,352	71	403	53	335	99	6,408	223
Total	9,063	1,503	1,586	385	2,552	801	31,084	2,689

^a Simple assault includes intimidation

^b Unknown

^c Stratum total (*N*) includes both unfounded and general incident reports (31,084)

were manually pulled and read for accuracy. In reviewing each record, definitions provided by the UCR program officials at the FBI were applied.⁶

A systematic procedure for the assessment of each record was established to ensure a high level of reliability between reviewers.⁷ For each record, a system of verifying the inter-rater agreement among reviewers was constructed by having multiple reviewers examine those records with a suspected classification error. When a classification error was suspected in a given record, at least three members of the research team (including the first author) reviewed the report to either confirm the classification error or reaffirm the accuracy of the original classification. If it was difficult to determine whether the report was classified correctly, the record was judged not to contain a classification error. Essentially, this procedure gave the reporting officer the benefit of any doubt in the absence of clearer information; thus, the research team relied on the judgment of the officer on the scene when crime classification was difficult to determine because the narrative was vague. This rarely happened in the current study.⁸

⁶ We have found in our experiences so far that police officers tend to write reports mostly in order to provide a full enough accounting of the incident in order to serve as an official memory of events for court testimony or if called before an administrative board. Therefore, we have had very few cases where we did not have enough information to classify the crime according to UCR definitions.

⁷ Prior to reviewing the records, a full, one-day training was provided by an official FBI trainer which focused on the proper classification of crimes in the UCR. The FBI provided materials to assist in this process.

⁸ Less than 5 records out of 2, 663 resulted in this type of resolution.

Table 4 Error estimate and sample size by offense category

	Total reports (<i>N</i>)	Estimated error (%)	Sample (<i>n</i>)
Arson	137	– ^c	59
Aggravated assault	673	0.08	210
Simple assault ^a	4,480	0.08	305
Burglary	2,522	0.15	444
Murder	15	–	9
Other homicide	0	N/A	0
Larceny	8,428	0.22	678
M. V. thefts	1,056	0.05	164
Robbery	366	0.05	135
Rape	125	0.02	32
Other sex offenses	142	0.02	50
Other Group A	6,651	0.08	299
Unfounded	31	–	31
General INCIDENTS	50	–	50
Group B	6,408	0.05	223
Total	31,084 ^b		2,689

^a Simple assault includes intimidation

^b Twelve municipal police agencies with resident populations over 10,000 compose the population for this study

^c Error not estimated due to small *N*

Calculating Statistical Accuracy

As stated previously, the focus of this study is on the statistical accuracy of crimes reported in the 12 largest municipal police departments in a southeastern state in the United States. The analysis begins with a presentation of results based on an assessment of agency records. These results are followed by an assessment of common classification errors found in individual crime categories and the associated impact of these errors on the aggregate crime totals. Based on a review of police records, this discussion is followed by a qualitative description of why many of the errors in classification occurred.

The results shown in Table 5 are displayed in a matrix of overcounts and undercounts. The rows reflect the initial classification of the reports by the law enforcement officer, and the columns show the classifications based on the reviewers assessment using UCR crime reporting definitions. Two numbers are presented in the matrix cell; the top number is the sample and the bottom number is the population estimate.⁹ Following along the top row of Table 5, notice that there is a “1” under “other Group A” offenses. This indicates that 1 incident was originally reported as arson, but was assessed by the reviewers to be an Other Group A offense rather than an arson. Below the “1” in this cell is a “2”. This indicates that in the population (i.e., all 12 municipal agencies combined) it is estimated that “2” arsons have actually been classified as other Group A offenses. Finally, the last column provides the overcounts in the sample and the population estimates of overcounts by

⁹ For example, under arson there are two numbers, 55 and 128. The 55 represents the number of sample reports out of 59 that were initially reported as arson and were confirmed to be arson. The number 128 is the estimate of the number of arsons in the population of reports that were recorded accurately by the police.

Table 5 Matrix of overcounts and undercounts by offense category

	Arson	Aggravated assault	Simple assault ^a	Burglary	Murder	Other homicide	Larceny	M. V. theft	Robbery	Rape	Other sex offense	Other Group A	Unfounded	General incidents	Group B	Sample size ^b	Overcounts
Arson	55	0	0	0	0	0	0	0	0	0	0	1	0	3	0	59	4
Aggravated assault	128	0	0	0	0	0	0	0	0	0	0	2	0	7	0	137	9
Simple assault	0	197	5	0	0	0	0	0	3	0	0	3	1	1	0	210	13
Burglary	0	631	16	0	0	0	0	0	10	0	0	10	3	3	0	673	42
Murder	0	18	268	0	0	0	0	0	4	0	0	0	0	2	2	294	26
Other homicide	0	274	4,085	0	0	0	0	0	61	0	0	0	0	30	30	4,480	395
Larceny	0	0	0	423	0	0	4	0	0	0	0	2	2	0	0	431	8
Robbery	0	0	0	2,475	0	0	23	0	0	0	0	12	12	0	0	2,522	47
Rape	0	0	0	0	9	0	0	0	0	0	0	0	0	0	0	9	0
Other sex offense	0	0	0	0	15	0	0	0	0	0	0	0	0	0	0	15	0
Other Group A	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Unfounded	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Arson	0	0	0	15	0	0	651	0	2	0	0	5	2	2	0	677	26
Aggravated assault	0	0	0	187	0	0	8,104	0	25	0	0	62	25	25	0	8,428	324
Simple assault	0	0	0	0	0	0	3	158	1	0	0	0	3	0	0	165	7
Burglary	0	0	0	0	0	0	19	1,012	6	0	0	0	19	0	0	1,056	44
Murder	0	0	0	0	0	0	1	0	134	0	0	0	1	0	0	136	2
Other homicide	0	0	0	0	0	0	3	360	0	0	0	0	3	0	0	366	6
Larceny	0	0	0	0	0	0	0	0	0	29	3	0	0	0	0	32	3
Robbery	0	0	0	0	0	0	0	0	0	113	12	0	0	0	0	125	12
Rape	0	0	0	0	0	0	0	0	0	8	38	0	3	1	0	50	12
Other sex offense	0	0	0	0	0	0	0	0	0	23	107	0	9	3	0	142	35
Other Group A	1	0	0	2	0	0	2	0	0	0	2	290	0	2	1	300	10
Unfounded	22	0	0	44	0	0	44	0	0	0	44	6,431	0	44	22	6,651	220
Arson	0	0	0	0	0	0	0	0	0	0	0	1	29	1	0	31	2
Aggravated assault	0	0	0	0	0	0	0	0	0	0	0	1	29	1	0	31	2
Simple assault	0	0	0	0	0	0	0	0	0	0	0	1	29	1	0	31	2

Table 5 continued

	Arson	Aggravated assault	Simple assault ^a	Burglary	Murder	Other homicide	Larceny	M. V. theft	Robbery	Rape	Other sex offense	Other Group A	Unfounded incidents	General incidents	Group B	Sample size ^b	Overcounts
General incidents	0	3	2	1	0	0	1	0	0	0	0	6	0	35	2	50	15
Group B	0	3	2	1	0	0	1	0	0	0	0	6	0	35	2	50	15
	0	0	3	0	0	0	1	1	0	0	0	0	0	0	214	219	5
	0	0	88	0	0	0	29	29	0	0	0	0	0	0	6,262	6,408	146
Total actual	56	218	278	441	9	0	663	159	144	37	43	308	41	47	219	2,663	-
	150	908	4,191	2,707	15	0	8,223	1,041	462	136	163	6,624	100	148	6,316	31,084	-
Undercounts	1	21	10	18	0	0	12	1	10	8	5	18	12	12	5	-	133
	22	277	106	232	0	0	119	29	102	23	56	93	71	113	54	-	1,297

^a Simple assault includes intimidation

^b Sample sizes differ slightly from Table 3 due to some records not being available for review

offense category. While the bottom row denotes the undercounts in the sample and the population estimates of undercounts for each offense.

As shown in the bottom right-hand cell of Table 5, there were 133 out of 2,663 records in the sample that contained classification errors (considering both overcounts and undercounts). Based on the error rate in the sample it was estimated that 1,297 classification errors were contained in the population of 31,084 records. Thus, approximately 4.17% of all records reported in the population of 12 agencies were estimated to be misclassified. We found that some crime types were more prone to error than others. The crime categories that had the highest error estimates include: simple assault/intimidation ($501 = 395$ estimated overcounts + 106 estimated undercounts), larceny ($443 = 324$ estimated overcounts + 119 estimated undercounts), aggravated assault ($319 = 42$ estimated overcounts + 277 estimated undercounts), other Group A ($313 = 220$ estimated overcounts + 93 estimated undercounts), burglary ($279 = 47$ estimated overcounts + 232 estimated undercounts), Group B ($200 = 146$ estimated overcounts + 54 estimated undercounts), and general incident ($128 = 15$ overcounts + 113 estimated undercounts). On the other hand, far less or no classification error was found in other crime types, such as Murder for which no misclassified records were found.

The crimes of motor vehicle theft, arson, and rape also contained classification error. In the case of arson, only 31 records were estimated to contain an error, 9 over counts and 22 undercounts. As a result, misclassifications associated with the offense of arson did not contribute a great deal to the overall level of classification error in UCR statistics. Similarly the crime of rape contained few classification errors. It is estimated that a total of 35 records in the population contained an error. Of these 35 instances, all of the misclassifications occurred between the offense of rape and other sex offenses. A total of 12 crimes were originally classified as rape and were subsequently deemed to be other sex offenses. Meanwhile, 23 crimes initially recorded as other sex offenses were assessed by the reviewers as rape cases. This resulted in a net reduction of 11 offenses that were actually rape, but classified otherwise.

In addition, a close examination of Table 5 reveals a general pattern for much of the classification error that occurred *between* individual crime types. For example, much of the error found in larceny tended to involve burglary and vice versa. Similarly, a great deal of the error associated with aggravated assaults was related to the error found in simple assault/intimidation.

In addition, an estimated 274 aggravated assault offenses were misclassified by law enforcement officers as simple assaults/intimidation. Far fewer aggravated assaults were misclassified as simple assaults. Only 16 offenses were initially recorded as aggravated assault and later assessed as simple assault. After the net gains and losses were considered, a total of 258 misclassifications occurred due the difficulty in differentiating between these two crimes. The end result was a considerable downgrading of aggravated assaults to the less serious offense of simple assault.

While there was considerable variability in the magnitude of error across crime types, a majority of misclassifications tended to occur in a predictable fashion. An exception to this is was found in the crime robbery. Where robbery is often thought to be reported correctly, we estimate that 102 robberies were undercounted. Sixty-one of these undercounts (60%) were found in the simple assault category while 25 were found in the larceny category. To some degree this is understandable since robbery is defined as theft by force or threat of force. The following section illustrates the impact of the classification error found among individual crimes on the total population of reported offenses.

Table 6 Crime estimates, error rates, and confidence intervals by offense category

	Estimate	Reported	Ratio	Error (%)	Low	High
Arson	150	137	0.91333	-8.67	107	194
Aggravated assault	908	673	0.74199	-25.88*	788	1,028
Simple assault ^a	4,191	4,480	1.06896	6.90*	4,019	4,363
Burglary	2,707	2,522	0.93166	-6.83*	2,595	2,819
Murder	15	15	1.00000	0.00	-	-
Other homicide	0	0	-	-	-	-
Larceny	8,223	8,428	1.02493	2.49*	8,077	8,369
M. V. thefts	1,041	1,056	1.01441	1.44	977	1,105
Robbery	462	366	0.79221	-20.78*	530	934
Rape	136	125	0.91911	-8.09	120	152
Other sex offenses	163	142	0.87117	-12.88	101	227
Other Group A	6,524	6,651	1.01947	1.95	6,379	6,665
Unfounded	100	31	0.31000	-69.00*	57	143
General incidents	148	50	0.33784	-66.22*	68	228
Group B	6,316	6,408	1.01457	1.46	6,178	6,454
Total	31,084	31,084	1.00000	0.00	-	-
Violent crime total ^b	1,521	1,179	0.77515	-22.49*	1,382	1,660
Property crime total ^c	12,121	12,143	1.00182	0.18	11,923	12,319
Index total	13,642	13,322	0.97654	-2.35*	13,399	13,885

* Statistical significance at $p < 0.05$ level

^a Simple assault includes intimidation

^b Includes aggravated assault, murder, other homicide, robbery, and rape

^c Includes arson, burglary, larceny, and motor vehicle theft

Findings

Tables 6 and 7 provide a summary of point estimates, confidence intervals, and statistical error rates for each reported crime category. Both tables also offer an estimate of the total number of Index crimes,¹⁰ violent, and property crimes. The “Estimate” column is the point estimate of the true number of crimes in each category and is calculated according to Eq. 2. The “Reported” column shows the original number of crimes (N) reported in each category. The “Ratio” column is a comparison of the reported numbers to the true numbers with the true number as the reference variable in the denominator, see Eq. 3. The “Error” column is the percent overcount or undercount in each crime category and is calculated according to Eq. 3. Negative percent error indicates an undercount of crimes in a given category; positive percent error is indicative of overcounts in a given crime type. The “low” column in each table reports the lower bounds of the confidence interval, and the “high” column represents the upper bounds. The confidence intervals were calculated according to Eq. 5.

¹⁰ Index crimes are murder, rape, robbery, aggravated assault, burglary, motor vehicle theft, larceny, and arson.

Table 7 Crime estimates, error rates, and confidence intervals with simple assault considered an index crime by offense category

	Estimate	Reported	Ratio	Error (%)	Low	High
Arson	150	137	0.91333	-8.67	107	194
Aggravated assault	908	673	0.74199	-25.88*	788	1,028
Simple assault ^a	4,191	4,480	1.06896	6.90*	4,019	4,363
Burglary	2,707	2,522	0.93166	-6.83*	2,595	2,819
Murder	15	15	1.00000	0.00	-	-
Other homicide	0	0	-	-	-	-
Larceny	8,223	8,428	1.02493	2.49*	8,077	8,369
M. V. thefts	1,041	1,056	1.01441	1.44	977	1,105
Robbery	462	366	0.79221	-20.78*	530	934
Rape	136	125	0.91911	-8.09	120	152
Other sex offenses	163	142	0.87117	-12.88	101	227
Other Group A	6,524	6,651	1.01947	1.95	6,379	6,665
Unfounded	100	31	0.31000	-69.00*	57	143
General incidents	148	50	0.33784	-66.22*	68	228
Group B	6,316	6,408	1.01457	1.46	6178	6,454
Total	31,084	31,084	1.00000	0.00	-	-
Violent crime total ^b	5,712	5,659	0.9907	0.93	5,491	5,933
Property crime total ^c	12,121	12,143	1.0018	0.18	11,924	12,319
Index total	17,833	17,802	0.9983	-0.17	17,536	18,130

* Statistical significance at $p < 0.05$ level

^a Simple assault includes intimidation

^b Includes aggravated assault, murder, other homicide, robbery, and rape

^c Includes arson, burglary, larceny, and motor vehicle theft

$$T_i \pm 1.96 \sqrt{\frac{P_i Q_i N_i - n_i}{n_i - 1} \frac{1}{N_i}}, \quad (5)$$

where T_i , the point estimate of total number of crimes for crime category i ; P_i , percent error found in the sample; $Q_i = 1 - P_i$; n_i , sample size; N_i , total reports in the population of crime category i .

All of the results in Tables 6 and 7 are identical, with the exception of the information used to calculate the Violent Crime and Index totals. Table 7 includes simple assault/intimidation as an Index crime. The results and implications for the inclusion of simple assault/intimidation as an Index crime are discussed later. But first, the results displayed in Table 6 illustrate the amount of classification error estimated for each crime type in the population. As shown in the error column, the highest "percent error" was observed in unfounded (-69.00%), general incident (-66.22%), aggravated assault (-25.88%) and robbery (-20.78%) offenses. All of the error is negative, indicating that these crimes were undercounted in original crime reports.

To a lesser degree, error percentages were reported for burglary (-6.83%), other sex offenses (-12.88%), arson (-8.67%), and rape (-8.09%). Similar to the other crimes, these reports were undercounted in original UCR statistics. It is important to note, however, that large error rates do not automatically translate into statistical significance. To

assess whether the observed error rates in each crime category were statistically significant, a 95% confidence interval was calculated. If the number originally reported by police did not fall within this confidence interval it was identified as statistically significant at $p < 0.05$.

The index offenses that were found to have statistically significant error estimates include aggravated assault, burglary, larceny, and robbery (see Table 6). Using robbery as an example, a total of 366 records were reported by law enforcement during the year under study. Based on a review of a sample of records, we estimated that 462 robberies should have been reported. Our confidence interval ranged from 530 to 934 so we conclude the classification error in the crime of robbery is statistically significant in this study.

In addition to index offenses, the error in the number of unfounded and general incident cases reported by law enforcement was significant. A total of 31 unfounded and 50 general incident records were reported for the population in this study. Following the review, the actual number of records was estimated to be 100 unfounded and 148 general incidents cases. Since the reported number did not fit within the confidence intervals these were also statistically significant.

The Impact of Classification Error on Aggregate Crime Totals

The violent crime total comprises murder, aggravated assault, robbery, and rape. The property crime total consisted of arson, burglary, larceny, and motor vehicle theft. All of the offense categories that make up the violent crime and property crime totals constituted the Index Total. As shown in Table 6, the violent crime total was undercounted by 22.49%. There were 1,179 violent crimes reported the law enforcement agencies that composed the population for this study. However, the estimated true number of violent crime offenses was 1,521. This undercount in violent crimes was statistically significant at $p < 0.05$. Similar to violent crimes, the undercount in the index total category was also statistically significant. A total of 13,322 index crimes were reported by law enforcement; however, the estimated number was 13,642.

Of the index offenses only the aggregate measure for property crime failed to reach statistical significance. In the case of property crimes, the numbers reported (12,143) and estimated (12,121) were nearly equal. This is interesting because two of the four crime categories that compose the Property Crime Total (burglary and larceny) had statistically significant amounts of error.

Based on all of the findings presented in Table 6, it appeared that most of the error in the violent crime and index crime categories was the result of misclassifications of aggravated and simple assaults. This was evidenced by the fact that aggravated assault was undercounted by 25.88%, while simple assault/intimidation was overcounted by 6.90%. To explore this situation more in-depth, simple assault/intimidation was added to the violent crime and index crime categories in Table 7.

As shown in Table 7, once all reported assaults (aggravated and simple) were included in the calculation of the violent crime and index crime totals, the amount of classification error was reduced considerably. The index total was now undercounted by only 0.17% and the violent crime total was also undercounted by less than one percent (0.93%). By combining the types of assaults in this manner we found that the error was no longer statistically significant.

Explanations for the Misclassification of Crime Types

Table 8 provides explanations for the misclassification in the five categories with the most errors: aggravated assault, simple assault, burglary, robbery, and larceny. Qualitative information gathered from the narrative of the police reports was used to identify specifically how the misclassifications occurred.

The first column in Table 8 reveals why misclassifications occurred in the aggravated assault category. A total of five different crimes were originally reported as aggravated assault, but later assessed by the research team to actually be another crime. These include simple assault, robbery, other Group A, unfounded, and general incident. As stated previously, the classification errors in aggravated assaults occurred mostly in relation to simple assault. For a crime to be classified as aggravated assault, an offender must have used a weapon or displayed the weapon in a threatening manner, or the victim must have suffered severe or aggravated bodily injury. In the absence of a weapon or such injury, a crime is typically classified as a simple assault. When misclassifications occurred in this study, they were mostly due to either *no mention of a weapon* being involved in the incident or the *absence of severe bodily injury*. Instead of serious injury, these reports tended to describe such injuries as a “knot on the head,” “redness to eyes,” “bruises,” and other “minor abrasions.” No weapons were reported as being involved in these incidents. Other aggravated assault misclassifications were related to robbery and Other A offenses such as destruction of property. In the case of robbery, police reports described acts of “shoplifting with a knife” or breaking “into a home with a gun and [taking] a DVD player”. On the other hand, other police reports described incidents that were more closely related to destruction of property. For instance, two police records noted that a “dog was shot” or “windows were shot out” with no mention of bodily injury to the victims nor did it mention that the victims were in danger of being injured—such as they were in the house at the time the damage was done to the windows.

The second column of Table 8 shows that some simple assaults were reported as aggravated assaults. Of the 26 overcounts for simple assault, more than two-thirds of these errors were more appropriately classified as aggravated assault. Most of these crimes, as reported by law enforcement officers, either involved a weapon or described severe bodily injuries to the victim(s). Various reports described the use of a weapon such as a beer or glass bottle or club. Other non weapon examples involved incidents in which a victim suffered “a broken nose and slight concussion” or statements noting severe injury such as “lacerated the victim’s finger.” Aggravated assault was the correct classification for these crimes due to the presence of a weapon(s) and/or severe injury to the victim(s).

To a lesser extent, some crimes were classified as simple assaults when they were in fact robbery. Robbery is defined in the UCR as the taking, or attempting to take, anything of value (US Department of Justice 2000) and can either be through force or threat of force or violence and/or putting the victim in fear of immediate harm or without these last elements. In most instances where a robbery was mistakenly classified as a simple assault, police records described acts in which force was used to take or attempt to take something of value from another person. For instance, police reports of simple assaults sometimes included statements such as “took cigarettes and beat up person,” “stole items from store and beat up store employee,” and “shoplifted and then attempted to strike employee with his car”. Since these incidents involved the taking of something by force and/or by putting the victim in fear of immediate harm, they were most appropriately classified as robbery.

The crime of robbery itself was over counted in this study primarily because of misclassifications in larceny, aggravated assault, or unfounded. In one case of robbery, the

Table 8 Summary description of common classification errors by offense type

Reported crime	Aggravated Assault	Simple assault	Robbery	Burglary	Larceny
Amended crime classification and explanation for overcount	<p>Simple assault Minor injuries and no weapon involved</p> <p>Robbery Threats with weapon and items stolen</p> <p>Other A Destruction of property</p> <p>Unfounded Self-defense</p> <p>General incident Not enough information to complete report or confirm aggravated assault</p>	<p>Aggravated assault Weapon involved with minor injuries</p> <p>Severe injuries with or without a weapon</p> <p>Robbery Weapon or threat of force to commit theft</p> <p>Group B Public intoxication</p> <p>Disorderly conduct</p> <p>Destruction of property (no on assaulted)</p> <p>General incident Civil disputes</p> <p>Suspicious activity</p>	<p>Larceny Grabbed money from person and no force or threat of force reported</p> <p>Unfounded Lost property, found after initial report</p> <p>Aggravate assault Weapon used in assault and/or serious physical injury but nothing taken or attempted to be taken from victim</p>	<p>Larceny Theft from vehicle</p> <p>Theft from associate</p> <p>Other A Destruction of property</p> <p>Entered building legally, deleted computer files</p> <p>Unfounded Misplaced property</p>	<p>Burglary Breaking and entering to commit a theft</p> <p>Robbery Weapon involved and/or threat of force to commit a theft</p> <p>Other A Warrant Forgery</p> <p>Destruction of property</p> <p>Unfounded Credit card malfunction, no crime</p> <p>Items missing, not stolen</p> <p>General incident Civil dispute</p>

The original crime classification by law enforcement officers is noted in the column headings. The section labeled "Amended Crime Classification and Explanation for Overcount" lists the reasons that crime classifications were changed based on UCR definitions. The short descriptions offer a brief explanation for why each of the five offenses were misclassified

narrative of the incident report indicated that the perpetrator “stabbed [a] guy with a razorblade, very serious injury was inflicted, and the victim was life-flighted to hospital.” This incident, as recorded by the police officer, contained no indication that this act was in pursuit of taking, or attempting to take, anything of value. Therefore, this incident should have been classified as an aggravated assault.

Another major source of classification error in this study involved the relationship between burglary and larceny. In many instances, burglary was overcounted by the misclassification of larceny and vice versa. Simply, burglary involves the unlawful entry into a building or other structure with the intent to commit a felony or theft (US Department of Justice 2000). Since burglary often involves larceny, these crimes were found to be prone to misclassification. The cases that produced the highest over counting of burglary involved the taking of items from a car. The UCR definition for burglary pertains solely to entry into a building or other structure with the intent to commit a felony. According to the FBI, a building or other structure does not include a car or personal vehicle. Instead, these offenses are to be reported as Larceny/theft offenses or, more specifically, theft from a motor vehicle. Over half of all overcounts for burglary involved the misclassification of larceny offenses.

The reverse was also true. A large proportion of overcounts in larceny were due to the misclassification of burglary. Of the burglary offenses erroneously reported as larceny, most involved the simple breaking into a house, apartment, or storage garage. Officers’ reports included such statements as “broke into house through window and stole items,” “entered building after it was closed and stole items,” and “storage lock broken, entered building and stole \$2,000 worth of items.” Since these incidents involved illegal entry into a building to commit the larceny, they were more accurately classified as burglary.

In addition to the misclassification of larceny as burglary, other crimes such as robbery, other Group A’s, unfounded offenses, and general incidents were also counted as larceny. Larceny cases that should have been robbery included threats to “cut a victim’s throat” and a shop employee being “shoved, pushed, and punched” during a shoplifting incident. Some reported larcenies should have been Other Group A offenses because the details of these crimes made them more appropriate for forgery, embezzlement, and destruction of property categories.

Discussion and Conclusions

This study sought to shine a light on the issue of classification error in the UCR. It also demonstrated a method for assessing this type of error. In addition, we showed where classification error fit within the context of the total error structure of UCR. So, before we discuss our overall findings, we want to say two things in general about this study. First, we stressed throughout this paper that this was a statistical study with the population being the 12 largest police departments in a southeastern state. So in terms of point estimates and confidence intervals, we have calculated estimates for this population only. However, we believe the findings are indicative of general problems in the UCR crime classification throughout the United States. The UCR was created so that police officers had uniform categories for reporting crimes across jurisdictions. So it is the uniform definitions and policies of the UCR program—i.e., not social context—that creates the types of classification problems. Here are just a few examples of what we mean: In every jurisdiction throughout the country UCR definitions are different in some way from local criminal statutes, so officers are forced to deal with the translation of local criminal statutes to UCR

crime categories. Robbery, by UCR definition, is the combination of assault and larceny; this fact creates opportunities for classification error when one or the other of these two crimes is less apparent. And, the UCR definition of burglary includes the crime of larceny as an element of the offense; therefore, it is likely to be mistaken for the less serious crime (larceny) particularly when unlawful entry is questionable or cannot be determined. For these reasons among others we think that the *types* of UCR classification errors that surfaced in our study are likely to be similar to those found elsewhere.

Second, we acknowledge that context does matter in crime reporting, such as whether a low-level crime will be reported to police in a high crime area or whether police in high crime areas will document less serious crimes even if the victim does come forward. The volume and type of crime that makes it into the police database is clearly affected by social context. From this standpoint our study of city agencies in a southeastern state does not provide much contextual diversity and would not be widely generalizable *if* our focus was reporting decisions. However, as we pointed out in “[Literature Review](#)” of this paper, we do not address unreported crime here. Our starting place is in the police database. Therefore, we believe that the limited context for our study is of less importance in terms of the generalizability of the findings than it would be if we were dealing with other aspects of UCR error. With that said, let’s take a look at our findings.

Of the 31,084 records in the population, a total of 1,297 records (4.17%) were estimated to have been misclassified. Several crime categories were found to contain significant amounts of classification error, such as aggravated assault, burglary, larceny, and robbery. Larceny was *over counted* by police while aggravated assault, burglary and robbery were significantly *undercounted* in official UCR statistics. Likewise, three non index offenses contained a significant level of error. These included simple assault/intimidation, unfounded offenses, and general incidents. The classification error found was determined to have a profound impact on aggregate crime totals. Violent crime and Index crime were significantly undercounted in this study, -22.49 and -2.35% , respectively. Meanwhile, the property crime total was slightly over counted; however, this small error was not found to be statistically significant.

This study also identified patterns in the misclassification of crimes. For instance, much of the error associated with larceny occurred in relation to the crime of burglary and vice versa. A similar pattern emerged between the crimes of aggravated and simple assault. These findings suggest that, perhaps, police have difficulty in making the fine distinctions that are necessary for accurately classifying crimes that are conceptually close in nature. Thus, while officers were often correct in determining that an assault had occurred, they had difficulty in deciding whether it was aggravated or a simple. The qualitative analysis of the police reports also seems to support this notion. In addition, when simple assault and intimidation are added to aggravated assault in the violent crime and index totals, the error in these totals shrink to less than 1%, demonstrating the impact of the importance of the aggravated-simple assault issue to the accuracy of the aggregate statistics.

The findings also suggest that characteristics of crime reporting, which hitherto have been determined from less refined analyses, may become clearer using this study’s methodology. For example, Rosenfeld (2007) suggested that the rise in UCR in recent years is in part due to the upgrading of simple assaults in the context of domestic violence to aggravated assaults. Our study does not address domestic violence directly, but it does point to a general trend the opposite way, i.e., the tendency to downgrade these crimes. Similarly, the crime of robbery has been considered by criminologists as one of the more accurately and persistently reported crimes. However, we found that robbery had a high

propensity for misclassification, very close to the rate of aggravated assault which has always been considered vulnerable to misclassification.

In addition to examining the extent of classification error, this paper set out to establish a methodology for assessing statistical accuracy of UCR crime reports. While a great deal of attention has been given to the study of factors that may influence victim reporting and law enforcement recording of crimes, no research has systematically assessed the impact of classification error on the statistical accuracy of UCR estimates. The methodology introduced in this paper seems to offer a reasonable approach for assessing the “statistical accuracy” of UCR crime statistics. The importance of this study resides in the fact that it provides a tangible example of how to assess classification error at the state level, where multiple police agencies contribute to the overall state crime statistics. We believe that the methodology applied in this paper can be readily transported to other states which report UCR data.

This research also provides insight into ways to improve the accuracy of the UCR crime statistics. First, the results of this study may assist police and UCR administrators in developing training for police officers to classify crimes according to UCR definitions. When the quantitative and qualitative results of this study or future studies are used together, ample information is provided for both the identification of and explanation for classification error in crime statistics. Given the inherent difficulty in making the fine distinctions necessary for classifying crimes, it is hoped that these qualitative accounts of cases may help police and state UCR officials pinpoint where particular difficulties in interpretation may reside among officers. The results of this study could be drawn upon to improve law enforcement training through the use of scenarios that illustrate common errors. This information may further lead to efforts on the part of UCR officials and police agencies to work at better preparing officers at making more precise distinctions between crimes for UCR reporting purposes.

Finally, based on the results of this study, it is clear that classification error is present in UCR statistics and that it varies by type of crime. Future efforts may seek ways to *statistically adjust* crime data for greater accuracy based on the magnitude and variation of known error in individual crime types as well as aggregate totals. Given the widespread reliance on UCR data, it is vitally important to continue efforts designed to improve the accuracy of crime reporting.

Acknowledgments The authors would like to thank Theresa Lester, Jeri Kirby, and Carly Jira Warren for their work on the original conceptualization of this project and the DCJS staff who assisted in the data collection effort. A special thanks to Dr. Yoshio Akiyama, the principal architect of the methodology used in this research, and Kevin MacFarland of the FBI for providing training on the reporting of NIBRS data and the interpretation of offense definitions. This project was funded by a grant from the US Bureau of Justice Statistics (Grant # 2004-BJ-CX-K004) and in part by the ASA/BJIS Statistical Methodology Research Program, a joint program of the American Statistical Association’s Committee on Law and Justice Statistics and the Bureau of Justice Statistics. The viewpoints expressed herein are solely those of the authors and do not reflect the opinions of the US Department of Justice, the Office of Justice Programs, the Bureau of Justice Assistance, or the American Statistical Association.

References

- Akiyama Y, Propher SK (2005) Methods of data quality control: for uniform crime reporting programs. Criminal Justice Information Services Division, Federal Bureau of Investigation, Washington
- Bachman R (1993) Predicting the reporting of rape victimization: have rape reforms made a difference? *Crim Justice Behav* 20:247–251

- Black D (1974) Production of crime rates. *Am Sociol Rev* 35:733–748
- del Frate AA, Goryainov K (1994) Latent crime in Russia. United Nations Interregional Crime and Justice Research Institute, Rome
- Dodge L (1990) Department of Justice's effort to complete state crime statistics [Abstract]. Abstract obtained from National Criminal Justice Reference Service
- Federal Bureau of Investigation (2007) Crime in the United States, 2007. Federal Bureau of Investigation, Washington
- Forero J (2000) Precinct's rosy crime rate was a distortion, the police say. *The New York Times*, New York City
- Gove W, Hughes M, Geerken M (1985) Are uniform crime reports a valid indicator of index crimes? An affirmative answer with minor qualifications. *Criminology* 23:451–501
- Greenberg MS, Ruback RB (1992) After the crime: victim decision making. Plenum, New York
- Hart T, Rennison C (2003) Reporting crime to the police. Bureau of Justice Statistics, Washington
- Lynch JP, Jarvis JP (2008) Missing data and imputation in the Uniform Crime Reports and the effects on national estimates [Abstract]. *J Contemp Crim Justice* 24:69–85
- Maltz M (1999) Bridging the gaps in police crime data: a discussion paper from the BJS Fellows Program. Bureau of Justice Statistics, Washington
- Maltz MD, Targonski J (2002) A note on the use of county-level UCR data. *J Quant Criminol* 18:297–318
- McCoy CR, Matza M, Fazlollah M (1998) Statistical manipulation by police goes back decades. *Philadelphia Inquirer*, Philadelphia
- New York State Division of Criminal Justice Services (n.d.) Crime reporting. Albany
- Nolan J, Haas SM, Lester TK, Kirby J, Jira C (2006) Establishing the "statistical accuracy" for Uniform Crime Reports (UCR) in West Virginia. West Virginia Criminal Justice Statistical Analysis Center, Charleston
- Ombudsmans Victoria (2009) Crime statistics and police numbers. Victorian Government Printer, Victoria
- Rashbaum WK (2003) West side crime statistics were softened, police say. *The New York Times*, New York City
- Rosenfeld R (2007) Explaining the divergence between UCR and NCVS aggravated assault trends. In: Lynch J, Addington L (eds) *Understanding crime statistics: revisiting the divergence of the NCVS and UCR*. Cambridge University Press, New York
- Schwind H, Zwenger G (1992) The dark number analysis of motives for nonreporting theft in three studies. *Stud Crime Crime Prev* 1:115–126
- Shah R, Pease K (1992) Crime, race and reporting to the police. *Howard J Crim Justice* 311:92–199
- Skogan W (1976) Crime and crime rates. In: Skogan W (ed) *Sample surveys of victims of crime*. Ballinger, Cambridge, pp 105–120
- US Department of Justice (2000) *Uniform Crime Reporting Handbook*, NIBRS edn. Federal Bureau of Investigation, Washington, DC
- Warner BD, Pierce GL (1993) Reexamining social disorganization theory using calls to the police as a measure of crime. *Criminology* 31(4):493–517
- Wexler C, Marx GT (1986) When law and order works: Boston's innovative approach to the problem of racial violence. *Crime Delinquency* 32:205–223
- White BB, Mosher DL (1986) Experimental validation of a model for predicting the reporting of rape. *Sex Coerc Assault* 1:43–56
- Woodard PL (1992) Assessing completeness and accuracy of criminal history record systems: audit guide [Abstract]. National Institute of Justice, Rockville