The NYPD's Compstat: compare statistics or compose statistics?

John A. Eterno† and Eli B. Silverman‡

†(Corresponding author) Department of Criminal Justice, Molloy College, 1000 Hempstead Avenue, PO Box 5002, New York, NY 11571-5002, USA. Tel:+1 1516 678 5000 ext. 6135; Fax: +1 516 256 2289; email jeterno@molloy.edu
‡John Jay College of Criminal Justice, 899 Tenth Avenue, New York, NY 10019, USA. Tel: 845 357 0685; email: Estcompany@optonline.net

John A. Eterno PhD is chairperson and associate dean of graduate studies in criminal justice at Molloy College. He is a retired captain from the NYPD. He is managing editor of Police Practice and Research: An International Journal. His books include Policing within the Law: A Case Study of the New York City Police Department, and Police Practices in Global Perspective (with Dilip Das). His peer reviewed articles are in various publications such as Professional Issues in Criminal Justice, International Journal of Police Science and Management, Women and Criminal Justice, and Criminal Law Bulletin.

Eli B. Silverman PhD is Professor Emeritus at John Jay College of Criminal Justice and the Graduate Center of CUNY. He has previously served with the US Department of Justice and the National Academy of Public Administration in Washington DC and was Visiting Exchange Professor at the Police Staff College in Bramshill, England. He has lectured and consulted with numerous police agencies in the US, Europe, Central and South Americas, Asia and Australia. His recent publications include: NYPD Battles Crime: Innovative Strategies in Policing, and peer reviewed articles featured in the International Journal of Police Science and Management and the British Journal of Community Justice.

ABSTRACT

The New York City Police Department (NYPD) introduced Compstat in 1994. Since then, the NYPD and many other law enforcement agencies which have adopted Compstat, have declared dramatic decreases in reported index crime. Indeed, in New York City the claim is that index crime has decreased over 76 per cent. Using both quantitative and qualitative methods, this research examines Compstat's managerial environment. Our main concern in this paper is examining the extent (if any) of pressures which managers believe they are exposed to at Compstat, as well as how those pressures might influence unethical crime reporting. Results indicate that managers perceive that they are subject to significantly greater pressure in the Compstat era especially, but not exclusively, with respect to decreasing index crime. Further, these pressures help explain unethical crime reporting in the Compstat era.

INTRODUCTION

Rarely has a development in policing evoked such a wide range of attention and opinion in both the academic literature and popular press. Compstat (short for compare statistics) is the New York City Police

Department's (NYPD) central management accountability model whereby commanders are held responsible for crime in their areas. Compstat features up-to-date computerised crime data, crime analysis and advanced crime mapping as the basis for regularised interactive crime strategy meetings which hold police managers accountable for specific crime strategies and solutions.

Initial assessments portrayed Compstat as an effective managerial crime-reduction accountability device (Bratton, 1997; Bratton & Knobler, 1998; DeLorenzi, Shane, & Amendola, 2006; Geoghehgan, 2006; Gurwitt, 1998; Kelling & Sousa, 2001; MacDonald, 2002; Maple & Mitchell 2000; Moore, 2003; O'Connell & Straub, 1999; Safir, 1998; Sanow, 2003; Schick, 2004; Silverman, 2001; Walsh, 2001; Walsh & Vito, 2004).

Some more recent Compstat appraisals, however, have offered significant reservations regarding Compstat's managerial effectiveness, the reliability of its crime statistics and/or the extent of its organisational reform (Dabney, 2010; Eck & Maguire, 2000; Eterno & Silverman, 2006; Harcourt, 2001; Harcourt & Ludwig 2006; Joanes, 2006; Karmen, 2000; Levitt, 2004; Messner et al., 2007; Rosenfeld, 2007; Rosenfeld, Fornango, & Baumer, 2005; Rosenfeld, Fornango, & Rengifo, 2007; Skogan & Frydel, 2003; Weisburd, Mastrofski, McNally, Greenspan, & Willis, 2003; Weisburd, Willis, & Greenspan 2006). Some critics claim that police middle managers are under substantial organisational pressure to contain crime in their areas. These detractors maintain that these pressures are reflected in an overemphasis on numbers and statistical performance management (Martin, 2003).

In New York City, these critical appraisals have been primarily fuelled by anecdotal accounts. With this in mind, this paper offers the first systematic survey of commanders who have been active in the

NYPD Compstat process, the law enforcement agency which first introduced Compstat.

We sought to ascertain if: (1) Compstat was perceived by mid-level managers as adding pressure to lower their crime rates; and (2) if so, was this pressure perceived as contributing to inaccurate crime statistics? A questionnaire was specifically designed to elicit frank and candid understanding of the dynamics of middle-management operations and interface with higher NYPD echelons. This paper consists of four sections. To provide the context for our analysis, this paper first reviews literature relevant to the controversy surrounding Compstat. The second section discusses our research approach, while the third section provides an analysis of our findings. The concluding section discusses the limitations of this study and the implications of the findings for international crime statistics.

THE INTERNATIONAL COMPSTAT CONTROVERSY

Since Compstat has been adopted throughout the United States (within its first six years, more than a third of US cities with 100 or more officers claimed to have adopted Compstat [Weisburd et. al. 2001]) and, in many other countries, the issues raised by the controversy have international ramifications. This can be observed in New York, other American cities, and internationally beyond the US.

New York

Compstat's first three years corresponded with dramatic declines in the city's crime rate coupled with many accolades. The city's 12 per cent decline in 1994 (compared with a national drop of less than 2 per cent) grew to 16 per cent in 1995 and maintained 16 per cent in 1996. These decreases accounted for more than 60 per cent of the national decline during this

period (Silverman, 2001). Between the calendar years 1990 and 2008, according to the New York City Police Department (NYPD), major crimes declined over 76 per cent while homicide also declined almost 77 per cent.

On the other hand, it has been suggested that Compstat is now used by NYPD headquarters to bear down on the management of many street operations in order to lower the crime rate. Numbers, sometimes any numbers, rule the day. This ratcheted system, in the words of one participant, is 'wound up too tight'. A 20-year veteran Brooklyn detective put it this way, 'Compstat is everything. People are tired of being harassed, searched and frisked, and run off the streets. People are fed up; the cops are, too' (Marzulli & O'Shaughnessy, 2000, p. 17).

There have been numerous press reports critiquing New York City crime statistics. In 2003, New York police officials revealed that more than 200 felonies committed in the Manhattan's 10th precinct in the previous year were improperly downgraded as misdemeanours, making the crime rate in the area appear lower than it really was.[1] According to Rashbaum (2003, p. B5):

> Because the reported crimes were wrongly downgraded, the 10th precinct . . . recorded just 811 serious crimes at the end of 2002, compared with 876 in 2001, for a 7.42 per cent drop, according to police statistics. Without the changes, serious crime in the precinct would show an increase for 2002 of about 16 per cent, to more than 1,000 crimes, while the city over all had a modest decrease of 5.32 per cent, according to department statistics.
>
> [Since Compstat went into effect] at least five police commanders have been accused of reclassifying crimes to improve their statistics, which are reviewed

at sometimes contentious weekly Compstat meetings.

Two years later, a reporter analysed a 44 per cent increase in NYPD reports of lost property (a misdemeanour) between 1997 and 2004, while grand larcenies (a felony) experienced steep declines. The author suggests that this is an example of downgrading crime in order to reduce the rate of major crimes (Moses, 2005).

New York City's own Patrolmen's Benevolent Association presented a similar critique of Compstat (Zink, 2004):

> It was a great idea that has been corrupted by human nature. The Compstat program that made NYPD commanders accountable for controlling crime has degenerated into a situation where the police leadership presses subordinates to keep numbers low by any means necessary. The department's middle managers will do anything to avoid being dragged onto the carpet at the weekly Compstat meetings . . .
>
> So how do you fake a crime decrease? It's pretty simple. Don't file reports, misclassify crimes from felonies to misdemeanors, under-value the property lost to crime so it's not a felony, and report a series of crimes as a single event. A particularly insidious way to fudge the numbers is to make it difficult or impossible for people to report crimes — in other words, make the victims feel like criminals so they walk away just to spare themselves further pain and suffering.

As recently as February 2010, a New York police officer reported that his precinct 'recorded felonies as misdemeanors and refused to take complaints from victims — all in an effort to drive down the crime rate' (Parascondola, 2010). Following other related accounts of crime manipulation, New York's Citizen Crime Commission, 'a

civic group that usually backs the NYPD urged the department to release internal audits of crime statistics to assure the public that the numbers can be believed' (Parascandola, 2010).

Other American cities

This Compstat allure of crime reduction, through technological advancement, is reflected in an in-depth study of the Lowell Police Department's Compstat (Willis, Mastrofski, Weisburd, & Greenspan, 2003a, p. 11):

> What police department, however, would not want to adopt a program whose clear purpose is to reduce crime through the implementation of a well-defined set of technologies and procedures? The appeal of Compstat's crime fighting goal to the police increases the likelihood that it will endure.

Compstat crime reduction efficacy is frequently advocated by police administrators, several of whom moved from the NYPD to head other city police departments. Compstat's introduction in New Orleans, for example, corresponded with a reported decline in murders from 421 in 1994, diving 55 per cent in 1999 to 162. Minneapolis's version of Compstat, CODEFOR (Computer Optimised Deployment-focus on Results), has been credited for a double-digit decrease in homicides, aggravated assaults, robberies, burglaries and auto thefts between 1998 and 1999 (Anderson, 2001, p. 4). In 2000, Compstat was introduced in Baltimore by its new chief, a former NYPD deputy police commissioner. By the end of the year, the city experienced below 300 homicides for the first time in 20 years, accompanied by an overall crime drop of 25 per cent (Anderson, p. 4; Clines, 2001, p. 15; Weissenstein, 2003, p. 27). Between 1999 and 2001, Baltimore's overall violent crime declined 24 per cent, homicides

dropped 15 per cent, shootings fell 34 per cent, robberies dropped 28 per cent, rapes 20 per cent and assaults 21 per cent (Henry, 2002, p. 307). Philadelphia's former police commissioner, another former NYPD deputy police commissioner, attributed a decline in the city's crime to Compstat-driven policing.

On the other hand, concerns have been raised in many jurisdictions that Compstat has served as a catalyst to inaccurate law enforcement statistical measurement (Long & Silverman, 2005; Manning, 2001; Willis et al., 2003b). When this type of system becomes excessively supervised, whether within a highly centralised organisation or from external political or hierarchical organisations, the consequences can be alarming. Subordinate units will naturally concentrate on those items being measured. Or, as the saying goes, 'what gets measured gets done'. This can lead to crime statistics manipulation and/or downgrading, which has been reported in numerous locales including Philadelphia, Atlanta (Hart, 2004, p. 6), New Orleans (Ritea, 2003a, p. 1, 2003b, p. 1), New York (Gardiner & Levitt, 2003, p. 8; Parascandola & Levitt, 2004, p. 5) and Broward County, Florida (Hernandez, O'Boye, & O'Neill, 2004, p. 9). In Philadelphia, charges of altered crime reports emerged after the police department introduced Compstat. 'If a person was punched in the eye, it might have been written up as a hospital report, so it didn't reflect a crime had occurred', reported one Philadelphia police official (Webber & Robinson, 2003).

International examples

Internationally, Compstat-like performance management systems have also been subjected to controversial claims. In the United Kingdom, for example, commentators have cited the 'emergence of US style criminal justice policies in other industrial democracies' (Jones & Newburn, 1997, p. 123).

Like the NYPD Compstat model, the UK approach stresses police results and outputs through the setting of explicit and measurable crime level indicators (Clark & Newman, 1997). There have been numerous UK manifestations of police performance management criteria dating back to 1983 (Jones & Silverman, 1984).

In 1997 the UK government directed the police forces to achieve 'best value' and 'crime reduction targets' (McLaughlin, 2007, p. 184) through 'league tables, ranking and performance measures' (Loveday, 2005, p. 150).

Yet reviews suggest that crime statistical performance indicators might create 'perverse incentives [that] compromise local innovation, efficiency and accountability' (Flanagan, 2008, p. 21). The collection and recording of crime statistics was questioned in a report entitled 'Collection and Accuracy of Police Incident Data', commissioned by the Home Office in 1996, which stated, 'there appears to be some variation not only in the number and type of events being recorded by the police, but also in the way certain events are interpreted for statistical purposes' (Portas & Mason, 1996, p. 24). A Home Office (2000) report revealed significant under-recording of crime by the police.

Other evidence suggested that in some cases crimes were reclassified so that burglary was recorded as criminal damage or other type of theft (Loveday, 1996). In 1999, a Police Complaints Authority investigation into recording practices in one force found that over 9,000 crimes were absent from official crime figures, indicating that the force's crime recording policy 'was designed to have the effect of artificially reducing recorded crime to a more politically acceptable level' (Davies, 1999, p. 3; for a fuller discussion see Hallam, 2009).

In addition to academic and government analyses, the popular press has publicised distorted crime statistics. In October 2008, it was reported ('Violent crime soars', 2008) that:

> Some police forces have been under-recording the most serious violent crimes, the Home Office said today, as it released figures showing a 22% increase. The category includes serious assault, murder, attempted murder and manslaughter.
>
> Officials admitted the under-counting could have been going on for more than 10 years. They said 13 forces were asked to re-examine their figures after they discovered some serious assaults were being recorded in a lower category of offence.

Two months later, the head of the UK Statistics Authority accused the Home office of releasing 'selective' knife-crime figures in order to downplay the extent of knife stabbings (Booth, 2008). By April 2009, a respected Home Office adviser observed that the public had little confidence in the accuracy of the government's crime statistics (Whitehead, 2009).

An in-depth analysis of three UK police forces involved in Compstat-like performance management activity concluded that '. . . the conflicting priorities brought about by managerial dictum and the bureaucratic rules governing the recording of crime are to define crime down. It leads to a manipulation of data to provide pleasing results' (Hallam, 2009, III). Another UK study referred to 'repeated reports of the massaging of figures by the police' (Martin, 2003, p. 161).

In Australia, Compstat-like performance management systems were also modelled after New York's Compstat. Queensland's Police Service version of Compstat is known as Operational Performance Review (OPR). One study found the introduction of OPR 'associated with a significant decrease in the total number of reported

offences' (Mazerolle, Rombouts, & McBroom, 2007).

The country's largest police force, the New South Wales police, developed its Operations and Crime Review (OCR) management system after visits to NYPD's Compstat (Davis & Coleman, 2000). A previous commissioner stressed that OCR policing is data driven, 'uncompromising difficult and stressful' (Ryan, 2000 cited in Kennedy 2010).

Like elsewhere, Australian performance-based policing is controversial. On the one hand, for example, two Australian scholars' evaluation of the OCR found this process to be effective in reducing three of the four offence categories studied (Chilvers & Weatherburn, 2004). Yet a 2000 evaluation by an independent consulting group (Hay Group Consulting Consortium, 2000) found communication to be largely a one-way process with little feedback to commanders, 'reinforcing the culture of fear and punishment'. The following year, the deputy commissioner resigned after he announced that crime was falling when the Bureau of Crime Statistics said it was increasing (Kennedy, 2000, p. 27).

Australia's National Uniform Crime Statistics Committee reported variations in crime statistics due to '. . . the extent of unreported crime; inadequacies in offence definitions, counting rules and offence classifications; procedural differences such as the offences under which an offender may be charged; differences in the way statistics are compiled as a result of the lack of uniformity in systems used and non-compliance with the rules governing the collation of statistics' (Hallam, 2009, p. 55) Carach and Makkai (2002) found that recorded crime statistics in the State of Victoria varied depending on whether an evidential or prima facie approach was applied by officers involved in the crime recording process.

The current study focuses on the reliability of crime statistics in the law enforcement agency which has the longest-running and most widely publicised Compstat process. It has been over 15 years since the NYPD first introduced Compstat to law enforcement. We now turn to our research approach and findings.

RESEARCH APPROACH

In this study, we examine the managerial environment which shapes Compstat practice in the NYPD. We do this using both quantitative and qualitative methods. Essentially, quantitative methods are used to gauge the levels of perceived organisational pressure as well as possible adverse consequences of those pressures on the ranks of captain and above. This includes the ranks of deputy inspectors, inspectors, and deputy chiefs who may command precincts, detective, transit, housing and other operational units. Qualitative methods, on the other hand, are used to develop a theoretical understanding of the data collected and to present those data in an accurate and comprehensible way.

Given that the NYPD (like many organisations) may not necessarily be forthcoming with information regarding questionable aspects of its widely publicised style of policing, we required an innovative research approach.[2] The cornerstones of our methodology are a self-report, anonymous, mailed survey with retired members of the NYPD in the ranks of captain and above, as well as in-depth interviews with various ranks. Surveys, when properly designed, allow researchers to examine sensitive issues such as drug use by juveniles or sexual behaviour. One important aspect of a survey, then, is its construction. In order to ensure that the survey was properly constructed, we spent months in the development process with the assistance of other researchers and focus groups including

retired and current NYPD personnel. The focus groups were limited to four participants to allow for discussion and review of the questions. They added several questions, revised wording and helped test the instrument. All advised that the questions were understandable and the instrument was concise. Other researchers, including but not limited to Molloy College's Institutional Review Board, advised that the instrument and methods were excellent for the purposes of scientifically examining the NYPD's managerial process. Further, subjects were well protected by the research design and instrument. Ultimately, after numerous drafts, a finalised survey was constructed.[3] The survey is short and easily self-administered. It is only two pages (ie, four sides) long. This is an important feature in that lengthy questionnaires can adversely influence response rate (see Neuman, 2000).

The sampling design is very important as well. The population studied is NYPD retirees in the ranks of captain and above. Using retirees is an important element of the method since these managers do not fear retaliation and, therefore, have a level of comfort that enables them to respond fully. To protect subjects further, and just as crucial to the method, is the use of anonymity for those administered the survey. Anonymity is an essential feature to a study examining sensitive questions (eg, illegal and/or unethical behaviour). Research on the use of questionnaires indicates that anonymity is an indispensable tool helping address the possibility of social bias in answers. Specifically, anonymity helps to achieve valid responses (Babbie, 1989; Bradburn, 1983; Dillman, 1983; Neuman, 2000). Bradburn (p. 298), for example, writes, 'With sensitive questions or those associated with a high degree of social desirability, the more anonymous methods of administration appear to work somewhat better'. Neuman (p. 258) puts it quite succinctly, 'Survey

methods that permit greatest anonymity are better for threatening questions. Thus, more honest answers may come from mail . . . questionnaires . . .'. We therefore chose a self-administered, mail questionnaire design that permits maximum anonymity of subjects.

Keeping the questionnaire very short and maintaining anonymity are key features that help increase sample size and validity, respectively. In our best judgment, the need for these features is critical. However, their use necessarily causes some trade-offs. For example, certain questions must be eliminated or simply cannot be asked, either to save space or to ensure anonymity. Although one may question such trade-offs, they were not made haphazardly but based on carefully considered scientific reasoning.[4]

The instrument was mailed to all retirees in the ranks of captain and above, with a cover letter, in September 2008. The use of regular mail generally has the advantage of protecting respondents better than an internet or email design (eg, use of cookies, possible hacking of responses, tracking computer messages and so forth (see Hagan, 2006). The cover letter was signed by the principal researchers and the Captain's Endowment Association (CEA) president in order to secure the widest range of cooperation. In the letter we ask for the respondent's cooperation in helping us obtain accurate information on the NYPD. We also sent a reminder postcard in December. By February 2009, we had a final return of 491 completed questionnaires. This represents a 41per cent response rate. For retired New York City civil servants, this is considered a good return.

The sample consists of 166 (33.9 per cent) respondents who retired before 1995 (the first full year of Compstat in New York City — it began in early 1994) and 323 (66.1 per cent) who retired 1995 and after.[5] These percentages closely comport with the

known levels of retirees who are in those categories. Furthermore, there were 263 (54.1 per cent) captains, 98 (20.2 per cent) deputy inspectors, 63 (13 per cent) inspectors, 28 (5.8 per cent) deputy chiefs, and 34 (7 per cent) others.[6] These percentages also approximate the ranks of those retired.[7] These data increase our confidence that there is little, if any, non-response bias. Additionally, this provides an excellent sample for comparing two important groups — the first group is those who never worked under a Compstat management system and the second is those who did work during the Compstat era.

In addition to the self-report survey, the authors conducted in-depth interviews with 10 retired members of the NYPD. While these interviews are not a representative sampling of the entire police department, they do include various ranks from police officer through deputy chief: one police officer, one sergeant, three captains, two deputy inspectors, two inspectors and one deputy chief. All worked during the Compstat era. Two worked regularly with Compstat in headquarters. All were in good standing and remain so at this time. All retired after serving the department until eligible to retire and none had any serious disciplinary problems. These members volunteered to assist us with the research by contacting us or advising us they would help.

The data for the qualitative portion of the study were gathered through unstructured interviews. These unstructured interviews are similar to what is called 'informal conversational interview'. As Maxfield and Babbie (2009, p. 205) state, '. . . the interviewer establishes a general direction for the conversations and pursues specific topics raised by the respondent. Ideally, the respondent does most of the talking'. This was the case with us, as respondents were full of information that they wanted to share. The interviews generally lasted about

an hour to an hour-and-a-half. They were casual and took place either through a conference call or in an office atmosphere.

Both researchers were present for every interview. There are advantages to both researchers working in tandem. First, we were both present to ensure that we both fully understood. If not, either could clarify what was being said. Second, both of us are witnesses to what is being said. Therefore, we can corroborate the statements of respondents. Third, this enhances the reliability of the data in that two separate researchers agree with the information collected. Further, after each interview, the researchers compared notes to be sure that the substance of the interview was properly recorded. Last, if necessary, follow-up calls can be done by either researcher to clarify information that may have been unclear.

Both authors have also been extensively involved in other research on NYPD and Compstat (see, eg, Eterno, 2001, 2003; Eterno & Silverman, 2006; Silverman, 2001; Silverman & O'Connell, 1998). Also, one of the principal researchers is a retired captain from the NYPD, has attended Compstat meetings, and has discussed these issues with many of his colleagues.[8] Additionally, the other principal researcher has written what some consider one of the preeminent explorations of Compstat. Further, he has attended numerous Compstat meetings and has interviewed many members of NYPD in conducting that research. Collectively, these experiences and studies inform our understanding of NYPD and these data.

ANALYSIS AND RESULTS

Throughout this paper, we develop our understanding of NYPD using both quantitative (primarily) and qualitative (supplementary) methods. In this section, we focus on quantitative analysis. However, interpretation of these data and results is also

informed by the qualitative research (personal interviews) and our collective understandings of NYPD and Compstat. Our research focus in this paper is on the extent to which managerial style, especially Compstat pressures and the like, influence (if at all) the accuracy of crime statistics.

We use several statistical approaches to analyse the survey data. First, we use a comparison of means, examining the responses of those who worked during the Compstat era versus those who did not. This gives us baseline measurements comparing pressure during pre- and post-Compstat eras. Next, we examine the responses of those who worked during the Compstat era with respect to their self-reported pressures and changing of crime reports. At a minimum, this will show the extent (if any) of adjustment of crime reports. Lastly, we conduct a multivariate ordinary least squares regression analysis to attempt to explain the unethical modification of crime reports that respondents reported during the Compstat era. We supplement these quantitative data with qualitative data to help understand the nature of the findings and guide the research.

Compstat era perceptions of pressure compared with non-Compstat era perceptions

The first analysis is a comparison of means for those who worked after 1994 compared with those who worked earlier years (ie, those who worked during Compstat compared with those who never worked under Compstat) (see Table 1).[9] If there are no statistical differences between these eras, then these variables are not likely to be relevant to our research questions.

Nearly every variable demonstrates higher levels of managerial pressure perceived by those who worked during the Compstat era. As noted in Table 1, this is particularly evident in such areas as reducing index crime, increasing numbers of

arrests and summonses, and even reducing civilian complaints. One exception to this is pressure to maintain integrity in crime statistics. For this variable, pressure is significantly greater in the pre-Compstat era. Further, the only variable that does not reach the level of statistical significance is pressure to find victims of child abuse.

For the purposes of this study, we focus on three key variables in the survey: (1) the amount of pressure to decrease index crime; (2) the level of NYPD promotions based on crime statistics; and (3) the extent of management's demand for integrity in crime statistics. These variables establish a foundation for further analysis by showing respondents' varying perceptions before and after Compstat. Further, comments by our respondents indicate that these are critical variables. As an example, one respondent states:

> As crime goes down, the pressure to maintain it got great . . . it was a numbers game . . . promotions were based on favouritism and nepotism . . .

At the outset, then, we determine the extent to which (if at all) Compstat increases pressure to decrease index crime compared with pre-Compstat management policies. Some studies suggest Compstat enormously increases the pressure on commanders to reduce crime (eg, Eterno, 2001; Eterno & Silverman, 2006; Silverman 2001). However, this may be subject to doubt, since fighting crime is clearly within the mission of all police departments of any era, with or without Compstat.

The survey questions analysed here are based on a 1 to 10 scale. For this first analysis, '1' corresponds to 'least pressure' to reduce crime and '10' corresponds to 'most pressure'. We find that, on average, retirees who worked after 1994 (the Compstat era) felt far more pressure to decrease index crime ($M = 8.26$, $SE = 0.116$) compared with those who worked in 1994 and

Table 1: *T*-tests comparing Compstat era to non–Compstat era NYPD managers

Variable	Compstat era (Yes or No)	N	M	SD	Mean difference
Increase summonses	No	141	5.48	2.55	−1.461★★★
	Yes	313	6.94	2.14	
Increase arrests	No	139	4.94	2.40	−1.839★★★
	Yes	315	6.78	2.11	
Decrease index crime	No	137	5.66	2.96	−2.607★★★
	Yes	315	8.26	2.05	
Decrease other crime	No	135	5.22	2.58	−0.935★★★
	Yes	312	6.16	2.42	
Downgrade index crime to non-index crime	No	133	2.51	2.32	−1.365★★★
	Yes	316	3.88	3.05	
Improve quality of life	No	139	5.97	2.61	−0.991★★★
	Yes	316	6.96	2.08	
Decrease CCRB complaints	No	140	5.21	2.83	−1.29★★★
	Yes	317	6.50	2.30	
Increase stop and frisk reports	No	137	3.57	2.19	−2.205★★★
	Yes	315	5.77	2.50	
Detect victims of domestic violence	No	139	3.57	2.34	−1.320★★★
	Yes	313	4.89	2.53	
Detect victims of rape	No	138	3.51	2.44	−0.792★★★
	Yes	313	4.31	2.53	
Detect victims of child abuse	No	137	4.03	2.66	−0.45
	Yes	313	4.48	2.65	
Level of promotion based on crime statistics	No	153	5.92	2.47	−0.93★★★
	Yes	309	6.84	2.54	
Extent of management's demand for integrity in crime statistics	No	158	7.18	2.35	0.655★★
	Yes	316	6.52	2.80	

Notes:
★★★*p*<0.001, ★★*p*<0.01. NB: 'Level of promotion based on crime statistics' is based on answer to question 7 and 'extent of management's demand for integrity in crime statistics' is based on answer to question 9. Other variables are from question 1.

before (M = 5.66, SE = 0.253). This difference is statistically significant: t (195.081) = −9.369, p<0.001.[10] We take the added step of calculating the size effect using the formula:

$$r = \sqrt{\frac{t^2}{t^2 + df}}$$

Our calculations indicate that these data represent a large size effect (r = 0.557).[11] Clearly, those working in the Compstat era felt much more pressure to reduce index crime compared with those who worked in earlier years.

Next, we examine the extent to which commanding officers felt that their promotions were based on crime statistics. On average, those retirees who worked after 1994 were more likely to feel that promotion was based on crime statistics (M = 6.84, SE = 0.144) compared with those who retired earlier (M = 5.92, SE = 0.199). This difference is statistically significant: t (460) = 3.739, p<0.001 and represents a small size effect (r = 0.172). This finding is consistent with the previous result; namely,

those working in the Compstat era feel more pressure to reduce crime, which is also reflected in their views of promotional opportunities.

Another important comparison of means examines management's demand for integrity in crime statistics. One might assume that integrity would be the same regardless of the era. Our data, however, indicate that, on average, those who worked *after* 1994 felt *less* demand for integrity in crime statistics ($M = 6.52$, $SE = 0.157$) compared with earlier retirees ($M = 7.18$, $SE = 0.187$). This difference is statistically significant $t (366.322) = 2.678$, $p<0.01$ and represents a small size effect ($r = 0.138$). So, according to this survey, demand for integrity in the crime statistics was significantly weaker during the Compstat era.

Overall, these bivariate findings clearly demonstrate that pressures from management are significantly different in the Compstat era. First, all but one variable show significantly more pressure on Compstat era managers. The one variable that did not show a statistically significant difference was detecting victims of child abuse. Second, the pressure to reduce index crime is particularly striking given the size effect of 0.557. Such a large size effect — rare in social science — indicates that perceived pressure to reduce index crime during the Compstat era is pronounced. Third, the one variable that shows significantly *less* pressure in the Compstat era is management's demand for integrity in crime statistics.

The significant Compstat pressures to reduce crime, stress crime statistics for promotion, and the weakened demand for integrity in crime statistics heighten the likelihood that Compstat era commanders might change or modify index crime reports. Consequently, we turn our attention to the possible distortion of crime statistics in the Compstat era. To do this, we first conduct some univariate analyses and

then conduct a multivariate regression analysis to attempt to explain unethical changes to crime reports.[12]

Analysis of Compstat era responses

Based on previous analysis and studies suggesting possible manipulation of crime reports due to Compstat (Eterno, 2003; Eterno & Silverman, 2006), we examine data specifically targeted to Compstat era managers regarding changing crime reports.[13] We asked respondents to advise if they were aware of any instances in which crime reports were changed *specifically due to Compstat*.

By design we asked this as a contingency question. This allows us to select out those who know of no changes at all. That is, if respondents were unaware of any changes due to Compstat, it is highly unlikely that they were in a position to observe much of anything since complaints are routinely changed due to Compstat. Many of these changes are completely ethical and appropriate. For example one interviewee, who was a CO of a precinct and we would describe as highly ethical and even very positive about the NYPD, routinely sent out a police officer to look for cars which were reported stolen but were actually parked in a spot that the complainant could not remember.

> I actively sent out an officer to fix the numbers. If a car was reported stolen but was actually parked in a spot that the person could not find the car, well that had to be corrected. It was a matter of accuracy to be reflected in the Compstat figures.

Thus, if a respondent was completely unaware of any changes, even ethical ones, that person clearly was not in a position (ie, work assignment) to observe much of anything with respect to crime numbers. The contingency question, therefore, has the

advantage of exposing those respondents who were in positions to observe changes — without having them reveal information about their assignments — helping the researcher to maintain anonymity (and, consequently, validity) while as the same time collecting important information.

Of those who worked in the Compstat era and responded to the question, over half (51 per cent) of the 309 responses indicated they were aware of changes. Since this was a contingency question, it is now possible (and important) to ask only those who were aware of at least one change, follow-up questions about the change(s). The follow-up question asks the extent to which the change(s) were ethically inappropriate. To examine this, we grouped responses into low (responses 1–3), medium (responses 4–7), and high (responses 8–10) unethical categories. Since it was a contingency question, there were 160 responses (eliminating those who saw no changes whatsoever — again, very unlikely if the respondent worked in a precinct for any length of time). Of those who observed at least one change in a crime report, over half of respondents (53.8 per cent) indicated that the changes observed were highly unethical. Another 23.8 per cent indicated they were aware of moderately unethical changes. The remainder, 22.5 per cent, indicated that they were aware of ethical changes.[14]

Our qualitative interviews provide further evidence of perceptions of unethical alterations of reports in the Compstat era. Most interviewees indicated that reports were routinely changed in various ways due to pressure from Compstat.

> One C.O. would check book and catalog to look up depreciated value to make it a misdemeanor, same with car parts. That's how desperate C.O.s would get. [They would] go to [the] scene where index crime occurred, assess, try to get [the] complainant to change story.

Here we see an example of what Commanding Officers (COs) believe was happening. One example has to do with the cost of stolen items. If a citizen registered a complaint of larceny of property, every attempt was made to determine that the cost of the item would not make the crime a felony (grand larceny is an index crime but petit larceny is not). Regardless of what the complainant stated, unless the complainant had absolute verifiable proof, officers would make every effort to ensure the crime did not meet the level of a felony. Indeed, some retirees advised that going on 'eBay' to find an item similar to the one reported stolen but at a lesser price (to make the complaint change to a petit larceny from a grand larceny) was a well-known practice. Interviewees advised of commanders carrying books that stipulate the depreciated value of items. Also, getting complainants to change stories can be seen in the above quote.

We see other examples over and over again. One interviewee stated:

> There was full adherence to penal law, very literal interpretations. They would surf the internet to find a bicycle price, question victims to ensure it was an index offense . . . COs were driven to reading the letter of the law.

Yet another:

> It was a matter of spiking [ie, a sharp increase in crime] then apply the letter of the law . . . try to find something to 'legitimately' knock down crimes.

Thus, the interviewees felt that every effort was made to ensure that a complaint of an index crime exactly conformed to the New York State Penal Law section's rules for that crime. Such practices are not technically wrong but they are certainly practices that impact interpretation of crime statistics.

Clearly, the evidence thus far suggests that members of the NYPD have, at times, unethically altered complaint reports (see also Eterno, 2003, p. 114; Eterno & Silverman, 2006). What would induce commanders to resort to such legerdemain? To explain further and add evidence to this issue, we conducted a multivariate analysis.

Multivariate regression to explain unethical changing of crime reports

We have established that a large subset of respondents who were active in the Compstat era reported being aware of crime reports being changed in an unethical way. Evidence for this is seen in our previous quantitative and qualitative data analyses. At this point in the analyses, it seems reasonable to propose that the substantial pressures we have seen to reduce crime are helping to explain these unethical adjustments. Based on the previous analyses, certain variables make better theoretical sense and should be included in the model. In particular, variables that include Compstat pressures, promotional fairness, and other similar variables seem most appropriate.

As a first step, Table 2 shows descriptive statistics for the variables in the multivariate model. Pressure to decrease index crime has the highest mean ($M = 8.79$, $SD = 1.655$). Given the previous analyses, this seems logical. The lowest mean (of variables measured on a 1 to 10 scale) is for fairness of the promotion system ($M = 3.58$, $SD = 2.091$). It should be noted that none of the respondents felt that the promotion system was exceedingly fair (none responded 9 or 10 on the 10-point scale).

A correlation matrix of all the variables in the multivariate model was also included (see Table 3). Multicollinearity does not appear to be an issue since the predictors do not correlate too highly with each other (ie, nothing greater than 0.9 — see Field, 2005, p. 186). Additionally, tests for the presence of multicollinearity were performed including variance inflation factor (VIF), indicating multicollinearity is not a problem (see Field, p. 197). As might be expected, downgrading index crime has the largest bivariate correlation with the dependent variable ($r = 0.323$). Again, downgrading is very different from decreasing index crime.

Table 2: Descriptive statistics of variables in multivariate model[a]

	N	Minimum	Maximum	Mean	Std. deviation
Decrease index crime	158	1	10	8.79	1.655
Decrease other crime	158	1	10	5.99	2.433
Downgrade index crime to non-index crime	159	1	10	5.52	3.056
Increase stop and frisk reports	159	1	10	6.53	2.349
Detect victims of rape	157	1	10	4.29	2.389
Promotion based on crime statistics within a managers command	155	1	10	7.14	2.542
Fairness of the promotion process above the rank of captain	159	1	8	3.58	2.091
Extent to which changes were ethically inappropriate	160	1	10	6.71	3.173
Valid N (listwise)	151				

Note:
[a] Selection of cases is based on respondent working during Compstat era and answering question 'extent to which change is ethically inappropriate'.

Table 3: Correlation matrix of variables in the multivariate model[a]

		Extent to which changes were ethically inappropriate	Decrease index crime	Decrease other crime	Downgrade index crime to non-index crime	Increase stop and frisk reports	Detect victims of rape	Promotion based on crime statistics within a managers command	Fairness of the promotion process above the rank of captain
Extent to which changes were ethically inappropriate	Pearson Correlation	1	0.193*	−0.045	0.323**	−0.066	0.102	0.090	−0.115
	Sig. (2-tailed)		0.015	0.571	0.000	0.405	0.203	0.266	0.148
	N	160	158	158	159	159	157	155	159
Decrease index crime	Pearson Correlation	0.193*	1	0.198*	0.319**	0.274**	0.131	0.103	−0.019
	Sig. (2-tailed)	0.015		0.013	0.000	0.000	0.102	0.202	0.818
	N	158	158	157	158	158	156	155	157
Decrease other crime	Pearson Correlation	−0.045	0.198*	1	0.113	0.016	0.281**	0.127	0.053
	Sig. (2-tailed)	0.571	0.013		0.156	0.840	0.000	0.117	0.510
	N	158	157	158	158	158	156	154	157
Downgrade index crime to non-index crime	Pearson Correlation	0.323**	0.319**	0.113	1	0.288**	0.030	0.143	−0.251**
	Sig. (2-tailed)	0.000	0.000	0.156		0.000	0.712	0.076	0.001
	N	159	158	158	159	159	157	155	158
Increase stop and frisk reports	Pearson Correlation	−0.066	0.274**	0.016	0.288**	1	0.337**	0.195*	−0.098
	Sig. (2-tailed)	0.405	0.000	0.840	0.000		0.000	0.015	0.220
	N	159	158	158	159	159	157	155	158
Detect victims of rape	Pearson Correlation	0.102	0.131	0.281**	0.030	0.337**	1	0.257**	0.134
	Sig. (2-tailed)	0.203	0.102	0.000	0.712	0.000		0.001	0.095
	N	157	156	156	157	157	157	153	156
Promotion based on crime statistics within a manager's command	Pearson Correlation	0.090	0.103	0.127	0.143	0.195*	0.257**	1	0.061
	Sig. (2-tailed)	0.266	0.202	0.117	0.076	0.015	0.001		0.453
	N	155	155	154	155	155	153	155	154
Fairness of the promotion process above the rank of captain	Pearson Correlation	−0.115	−0.019	0.053	−0.251**	−0.098	0.134	0.061	1
	Sig. (2-tailed)	0.148	0.818	0.510	0.001	0.220	0.095	0.453	
	N	159	157	157	158	158	156	154	159

Notes:
*Correlation is significant at the 0.05 level (2-tailed).
**Correlation is significant at the 0.01 level (2-tailed).
[a] Selection of cases is based on respondent working during Compstat era and answering question 'extent to which change is ethically inappropriate'.

Downgrading is a specific practice of removing a crime from an index category and placing it in a lower category (eg, grand larceny to lost property). Decreasing index crime is simply a general pressure to lower crime. Those who felt the pressure to downgrade were the ones most likely to change crime reports (at least in the bivariate model).

We now turn to the multivariate model. Regressing ethically inappropriate changing of crime reports onto the independent variables explained 22.8 per cent of the variance in the extent to which changes were ethically inappropriate (see Table 4). Five of the seven explanatory variables were statistically significant.[15] Note that all five of the significant variables are about perceived level of pressure from Compstat. Further, neither of the two variables on promotion were statistically significant predictors. This null finding was interesting since many respondents mentioned promotion as an important concern. We now turn to examining the statistically significant variables in order of the magnitude of their standardised beta values (largest to smallest). Recall that

these variables are measured on a scale of 1 to 10.

As would be predicted from the bivariate correlations, the largest beta coefficient in this model is associated with perceived pressure to downgrade index crime to non-index crime (*standardised* $\beta = 0.326$).[16] The large beta coefficient for this variable is an indication that it is one of the most important variables in the model. The unstandardised slope coefficient for this variable indicates that for each unit increase in perceived pressure to downgrade index crime, an increase of 0.335 in ethically inappropriate crime report changes is reported, when the effect of the other variables in the model is held constant. Thus, the pressure to downgrade index crime is a key variable explaining unethical distortion of crime reports — as pressure increases, unethical changing of reports also increases (a positive relationship).

The next largest beta coefficient for this model is associated with perceived pressure to increase stop and frisk reports (*standardised* $\beta = -0.315$). The unstandardised slope coefficient for this variable indicates that for each unit increase in pressure to write more

Table 4: Multivariate ordinary least squares regression; dependent variable, extent to which changes ethically inappropriate

	Unstandardised		Standardised
	B	Std. error	Beta
Constant	4.613	1.445	–
Decrease index crime	0.399	0.152	0.212★
Decrease other crime	−0.277	0.103	−0.213★★
Downgrade index crime to non–index	0.335	0.088	0.326★★★
Increase stop and frisk reports	−0.424	0.115	−0.315★★★
Detect victims of rape	0.275	0.110	0.211★
Promotion based on crime stats in manager's command	0.079	0.096	0.064
Fairness of promotion process above captain	−0.159	0.117	−0.107

Notes:
$R^2 = 0.228$.
★$p < 0.05$, ★★$p < 0.01$, ★★★$p < 0.001$.

stop and frisk reports, a decrease of 0.424 in ethically inappropriate crime report changes is reported, when holding constant the effect of the other variables in the model.

There are at least two possible explanations why an increase in pressure at Compstat to write more stop and frisk reports should lead to a decrease in unethical downgrading. Stop and frisk reports have been a very contentious area for the NYPD. There have been claims of racial profiling (see, eg, New York City, Civilian Complaint Review Board, 2001; Spitzer, 1999). As the pressures to write stop and frisk reports increase from managers at Compstat, it may be that commanders feel an informal message that 'anything goes'. Therefore, the ethics of adjusted crime reports may seem trivial in comparison. Unethical crime reporting may pale into insignificance compared with the practice of stop and frisk. This, in turn, may boost commanders' perceptions that crime report alteration is not unethical. Alternatively, it may be that commanders feel a sense of proactive policing. That is, those running Compstat will support their efforts at fighting crime and there is less need for adjustment of crime reports. Certainly, this variable requires further study and analysis.

The variable with the next largest beta coefficient is decrease other crime (*standardised* β = –0.213). The unstandardised slope coefficient for decreasing other crime indicates that for every unit increase in pressure to decrease other crime, a decrease of 0.277 in ethically inappropriate crime report changes occurs, holding the other variables in the model constant. This result also makes sense (especially since the coefficient indicates a negative relationship with the dependent variable). That is, although pressure to decrease other crimes (non index crimes) also exists, such pressure will make it less advantageous to downgrade major crimes into the other crime category.

Since index crime is usually the focus of Compstat, commanders will be less likely to rearrange crime reports if there is also pressure on them to reduce other crime. Based on these responses to the questionnaire, the in-depth interviews, and our collective understanding of Compstat, commanders feel more pressure to reduce index crime compared with reducing other crime. Indeed, commanders reported that the pressures to reduce index crime (M = 8.26, SE = 0.116) tends to be greater than the pressures to reduce other crimes (M = 6.16, SE = 0.137). So, whenever there is pressure to reduce 'other crimes', managers may be reluctant to alter reports unethically, since they might actually have to 'upgrade' or 'increase' crime classifications to remove them from the other crime category — something managers would not want to do given the other pressures on them that are more powerful (namely, reducing index crime).

The next largest beta coefficient is for perceived pressure to decrease index crime (*standardised* β = 0.212).[17] The unstandardised slope coefficient for pressure to decrease index crime indicates that for each unit increase in pressure to decrease index crime, an increase of 0.399 in ethically inappropriate crime report changes is seen, holding constant the effect of the other variables in the model. This variable is also intuitively pleasing in that it demonstrates that, as pressure from Compstat to reduce index crime increases, crime reports are more likely to be unethically changed (again, a positive relationship).

The last statistically significant variable is pressure to detect victims of rape (*standardised* β = 0.211). The unstandardised slope coefficient indicates that for each unit increase in pressure to detect victims of rape, an increase of 0.275 in ethically inappropriate crime report changes is seen, holding constant the effect of the other variables. This finding also makes sense in

terms of other findings. That is, since pressure to detect victims will ultimately lead to more index crime reports (since there are more victims), commanders will feel additional pressure to change crime reports unethically as pressure to detect victims increases.

Other variables in the model were not statistically significant. These include: promotion based on crime statistics within a manager's command, and fairness of the promotion process above the rank of captain. It is interesting to note that many managers mentioned the lack of fairness in the promotion process affecting their behaviours, yet this has no significant influence on adjusting crime reports.

DISCUSSION

Overall, based on these survey data and in-depth interviews with retirees, we can confidently state that Compstat era captains and above, feel significantly different intense pressure from high-level managers than those who did not work in the Compstat era. The pressure is particularly pronounced for decreasing the levels of index crime. The survey also demonstrates significantly lower perceptions of pressure for integrity in crime statistics during the Compstat period. Further, managers' promotions are considerably more likely to be based on crime statistics in the Compstat era. Importantly, many of these pressures are the main factors that explain the Compstat era respondents' high perception level of unethical distortion of crime reports.[18]

Are these Compstat problems inherent and inevitable? At this point it is difficult to say (see Eterno & Silverman, 2011 for a fuller exploration of this issue). Nevertheless, it does appear that our interviewees and others agree that the longer Compstat or any other police statistical management system is in place, the more readily it will slip into 'statistical adjustments'.

Numerous reports and studies of statistical alternations in Compstat-like performance management systems throughout the world (see our *Introduction* herein), strongly buttress our findings. Widespread and increasing political and organisational demands to meet or exceed numerical targets intensify pressures on crime reporting. (Butler, 2000)

These findings may very well compound what some already believe is the substantial and mounting evidence of statistical adjustments. Maltz (1999), in discussing the Uniform Crime Reports (UCR) used to provide data to publish the annual 'FBI Crime in the United States' report, indicates that the data provided to the FBI are lacking in quality. Since the reporting of data for many jurisdictions remains a voluntary activity and despite the efforts of the FBI to maintain data quality, the many gaps in the data make their use questionable. Feinberg (2000) is also critical of the manner in which crime statistics are gathered in the USA, suggesting that there are fundamental flaws in the crimes that form part of the UCR index. Furthermore, the temptation to distort crime statistics that the police themselves collect and collate should not be underestimated, especially if these statistics are a main criterion for 'success'.

We recognise, however, that our findings must be coupled with limitations. Some of the limitations pertain to very important features of our methodology — maintaining anonymity and keeping the questionnaire short. Admittedly, based on choices we made regarding which questions to include, we certainly could have captured other important variables. For example, we did not collect data on who served as commanders of precincts and we did not ask those in the non-Compstat era the level of unethical reporting. While there are several reasons for not asking, it is an admitted limitation that will have to be addressed by future research.[19] Of course, since this is a

case study, there are limitations with respect to generalisability (ie, external validity) as well. One possible issue is the question on crime reports that could be subject to interpretation (see Endnote 3). In contrast, however, our focus groups strongly suggested that the question is clear and concise, and they believe that respondents clearly understood that they were expected to base their answers to this on their personal experience.[20] Further study is certainly needed to clarify, confirm, expand and explain much of what we have done here. Regardless, we are confident that this study is a very significant contribution to the field.

In sum, this study is the first empirical examination in New York suggesting numerical adjustments. Based on these results, we feel that there is cause for concern in over-reliance on and interpretation of NYPD crime statistics. While additional study is necessary, we can preliminarily say, based on our findings, that crime statistics in Compstat-like jurisdictions, at a minimum, warrant careful scrutiny.

ACKNOWLEDGEMENT

NOTES

1. The NYPD created a Data Integrity Unit (DIU) after the Compstat process was ongoing. Additionally, it has had a Quality Assurance Division. Both the unit and division routinely check written complaint reports; however, neither fully addresses the concerns discussed herein.

2. One might contend that the NYPD's current lack of transparency is exemplified by the fact that visitors, including those from other law enforcement agencies, are generally no longer welcome at Compstat meetings. Further, the NYPD has resisted attempts by the City's own Commission to Combat Police Corruption to review the crime statistics.

3. The entire questionnaire is available upon request from the authors. Because of its importance to the study, we include the wording of question 4 here: 'Are you aware of any instances in which crime reports were changed due to Compstat?' If respondents answered yes, they completed follow-up questions.

4. For the purposes of our research, detailed questions about frequencies of changed crime reports were deemed unnecessary and, quite possibly, damaging to the research. For example, many might not respond for fear of possible retaliation as the questions probe deeply into individual behaviours. We are looking to examine the managerial style and simply require basic data. We have no intention of conducting a prosecutorial examination of members of the NYPD. Given significant research on the 'blue wall of silence' (see, for example, Rothwell & Baldwin, 2007), it was deemed absolutely necessary for respondents to have a sense of comfort which this wording afforded them.

5. Two respondents did not answer that question. To ensure anonymity, the exact year of retirement was not asked since the year, plus other information, would allow us to identify some respondents.

6. For example, some may have risen to higher ranks but maintained their membership in the CEA.

7. Data are not available. The results about the sample are based on discussion with the CEA.

8. Just to be thorough, we advise the reader that this researcher did not complete a survey.

9. Based on our data, it is possible that some who worked during the Compstat era are basing responses on the previous era (ie, telescoping into the non-Compstat era). However, the dramatic differences in responses by era, combined with our qualitative interviews and understandings of Compstat, strongly suggest the interpretations of this paper are accurate.

10. We use Levene's test for Equality of Variance. This test allows us to determine whether the assumption of homogeneity is violated if the significance value is less than 0.05.

11. See Field (2005, p. 32) for more information.

12. We also have a variable for 'legally inappropriate' changes to crime reports. Our statistical analyses with legally inappropriate changes are similar to ethically inappropriate (a bit less explanatory power in the multivariate models but generally similar). However, our interviews and understanding of NYPD suggest that there is more evidence to support ethically inappropriate changes to crime reports (as opposed to legally inappropriate changes). Therefore, we present the ethically inappropriate variable in our paper. A discussion of legally inappropriate can be found in our forthcoming book: Eterno and Silverman (2011).

13. Given the demonstrated significant differences in pressure (including significantly more pressure for integrity in crime statistics in the previous era), we think it very unlikely that crime reports in the previous era were manipulated to the extent we show in the Compstat era. Since we did not collect data on previous era manipulation of crime reports, we cannot compare eras for this variable. While it is possible that there is no change in the extent to which unethical manipulation takes place, the findings here — namely, the demonstrated differences in pressure based on the T-tests — strongly suggest that this is highly unlikely. Of course, further study on the issue is recommended. Regardless, empirically demonstrating differences in eras with respect to the types of pressures and the extent of pressures, as well as explaining unethical reporting in the Compstat era, are just some of our key contributions to the field in this paper.

14. Percentages may not add up to 100 due to rounding.

15. Explanatory power indicates that other variables not captured in the survey may help explain further, eg, other variables about Compstat, the police culture, etc. We ran many other regressions. We feel that this model best represents the data since it fits with our theoretical understanding of what is occurring, does not violate regression assumptions (based on various statistical tests and plots of residuals), and has good explanatory power compared with other models with this dependent variable. To keep the article manageable in terms of length, we do not present the numerous models and associated statistics; rather, as is typical of scientific papers, we present key information about our selected model and take the added step of offering the questionnaire for those interested in replication.

16. This variable had the highest beta coefficient in nearly all of our regression models for this dependent variable.

17. Again, pressure to decrease index crime is not the same as pressure to downgrade index crime to other crime categories. Downgrading is a specific practice of moving crimes to other

non-index categories. Decreasing index crime is a general perception about what high-level managers would like with respect to index crimes.

18. These pressures were perceived pressures and management may not have intended them. Regardless, this is the most accurate empirical evidence to date showing the level of pressure felt by managers and one of the major consequences thereof.

19. First, based on our understanding of NYPD, these were the NYPD managers who had intimate knowledge, in one way or another, of practices taking place there. Note that some may not have observed any changed reports due to their assignment; nevertheless, they witnessed the management style and are in an excellent position to comment on it. Second, we did have concerns that knowing specifically who commanders were, combined with other information, could lead anyone with access to the data to figure out who responded to the questionnaire — something we were not willing to compromise with. Regarding not asking those in the non-Compstat era about their reporting, our focus was on Compstat and its influence on managers. Therefore, that question was not deemed critical enough to keep in the short questionnaire — along with other questions.

20. Interestingly, to further the point and to reinforce our focus groups and our interpretation, one respondent wrote this unsolicited remark, 'In regards to Question #4, I think everyone in NYPD is "aware" of instances in which crime reports were changed due to Compstat. However, I marked "No" for my answer because the instances that I heard of could have been only

rumors and I have no factual information of such occurrences'. Further evidence that our interpretation is accurate is based on many respondents marking 'no' for their answer to this question (strongly buttressing our view). Beyond this, as stated earlier, the wording is by design to protect respondents. If anything, it is most likely that our method of collecting crime report changes *undercounts* since we allow multiple events. Further, respondents retired over a 15-year period and worked in numerous commands. This makes it highly unlikely that events are conflated.

REFERENCES

Anderson, D. C. (2001, winter). *Crime control by the numbers: Compstat yields new lessons for the police — and the replication of a good idea* (Ford Foundation Report). New York: Ford Foundation.

Babbie, E. (1989). *The practice of social research* (5th ed.). Belmont, CA: Wordsworth.

Booth, R. (2008, December 13). Home Office accused of releasing selective knife crime figures. *The Guardian*. p. 3.

Bradburn, N. M. (1983). Response effects. In P. Rossi, J. Wright, & A. Anderson (Eds.), *Handbook of survey research* (pp. 289–328). New York: Academic Press.

Bratton, W. J. (1997). Crime is down in New York city: blame the police. In W. J. Bratton, N. Dennis, R. Mallon, J. Orr, & C. Pollard (Eds.), *Zero tolerance: policing a free society.* London: IEA Health and Welfare Unit.

Bratton, W., & Knobler, P. (1998). *Turnaround.* New York: Random House.

Butler, T. (2000). Managing the future: a chief constable's view. In F. Leishman, B. Loveday, & S. Savage (Eds.), *Core issues in policing* (2nd ed.). Harlow: Longman.

Carach, C., & Makkai, T. (2002). *Review of Victoria police statistics* (Australian Institute of Criminology Research and Public Policy

Series No. 45). Canberra: Australian Institute of Criminology.

Chilvers, M., & Weatherburn, D. (2004). The New South Wales Compstat process: its impact on crime. *Australian and New Zealand Journal of Criminology.* Vol. 37, No 2, pages 22–48

Clark J. and Newman J. (1997) *The Managerial State*, London: Sage.

Clines, F. X. (2001, July 19). 'Police in Cincinnati Pullback in Wake of Riots' *The New York Times*, p. Al.

Dabney, D. (2010). Observations Regarding Key Operational Realities in a Compstat Model of Policing. *Justice Quarterly. 27*(1): 28–51.

Davies, N. (1999, March 8). Watching the detectives: how the police cheat in fight against crime. *The Guardian.* p. 3.

DeLorenzi, D., Shane, J. M., & Amendola, K. L. (2006). The Compstat process. *FBI Law Enforcement Bulletin, 73*(10), 34–38.

Dillman, D. A. (1983). Mail and other self-administered questionnaires. In P. Rossi, J. Wright, & A. Anderson (Eds.), *Handbook of survey research* (pp. 359–377). New York: Academic Press.

Eck, J., & Maguire, E. (2000). Have changes in policing reduced violent crime? In A. Blumstein, & J. Wallman (Eds.), *The crime drop in America*. Cambridge University Press.

Eterno, J. A. (2001). Zero tolerance policing in democracies: the dilemma of controlling crime without increasing police abuse of authority. *Police Practice and Research: An International Journal, 2*(3), 189–217.

Eterno, J. A. (2003). *Policing within the law: a case study of the New York City Police Department*. Westport, CT: Praeger.

Eterno, J. A., & Silverman, E. B. (2006). The New York City police department's Compstat: dream or nightmare? *International Journal of Police Science and Management, 8*(3), 218–231.

Eterno, J. A., & Silverman, E. B. (2011). *Unveiling Compstat: the global police revolution's naked truths.* London: Taylor & Francis (forthcoming)

Feinberg, G. (2000). Does greater fiscal investment in police services mean safer streets? *International Journal of Police Science and Management, 3*(2), 138–152.

Field, A. (2005). *Discovering statistics using SPSS* (2nd ed). London: Sage.

Flanagan, R. (2008). *The review of policing, final report*. Home Office: http://www.home office.gov.uk

Gardiner, S., & Levitt, L. (2003, June 21). NYPD: some crime stats misclassified. *New York Newsday*, p. 11.

Geoghehan, S. 'Compstat Revolutionizes Contemporary Policing', *Law and Order*, 54:4, 42–46.

Gurwitt, R. (1998, January). The comeback of the cops. *Governing*. P. 14–19

Hagan, F. E. (2006). *Research methods in criminal justice and criminology* (7th ed.). Boston: Allyn and Bacon.

Hallam, S. A. (2009). *Policing in the iron cage: the tensions between the bureaucratic mandate and street level reality.* Unpublished PhD dissertation, University of Bedfordshire.

Harcourt, B. E. (2001). *Illusions of order: the false promise of broken windows policing.* Cambridge, MA: Harvard University Press.

Harcourt, B. E., & Ludwig, J. (2006). Broken windows: new evidence for New York City and a five city social experiment. *University of Chicago Law Review, 73*(4) 271–289.

Hart, A. (2004. February 21). Report finds Atlanta police cut figures on crimes. *The New York Times*, p. 26.

Hay Group Consulting Consortium (2002) *Qualitative and Strategic Audit of the Reform Process (QSARP) of the NSW Police Service.* Police Integrity Commission: Sydney, Australia.

Henry, V. E. (2002). *The Compstat paradigm*. Flushing, NY: Looseleaf Law Publications.

Hernandez, J., O'Boye, S., & O'Neill, A. (2004, February 25). Sheriff's office scrutinizes its crime data. *SunSentinel*, p. 6.

Home Office. (2000). *Review of crime statistics: a discussion document*. London: author.

Joanes, A. (2006). Does the New York City Police Department deserve credit for the decline in New York City's homicide rate? *Columbia Journal of Law and Social Problems*, 33(2), 265–291.

Jones, T., & Newburn T. (1997). *Policing after the Act: Police and Magistrates' Courts Act 1994.* London: Policy Studies Institute.

Jones, S., & Silverman, E. (1984). What price efficiency. *Policing, 1*(1), 31–48.

Karmen, A. (2000). *New York murder mystery.* New York University Press.

Kelling, G. L. and Sousa, W. H. (2001). *Do Police Matter?* Civic Report No. 22. New York: Manhattan Institute.

Kennedy, M. (2000). *Zero tolerance policing and Arabic speaking young people.* Unpublished Masters Thesis, University of Western Sydney, Australia.

Kennedy, M. (2008). *The Police, Means, Ends and the Rule of Law.* Saarbrucken: VDM Publishing

Levitt, S. D. (2004). Understanding why crime fell in the 1990s. *Journal of Economic Perspectives, 18(1)*, 163–190.

Long, M., & Silverman, E. (2005). The Anglo–American measurement of police performance: Compstat and best value. *British Journal of Community Justice, 3*(3), 37–44.

Loveday, B. (1996). Crime at the core. In F. Leishman, B. Loveday, & S. Savage (Eds.), *Core issues in policing.* London: Longman. 6–28.

Loveday, B. (2000). Reforming the police from local service to State police. *Political Quarterly, 66*(2), 141–158.

Maltz, M. D. (1999). *Bridging Gaps in Police Crime Data: A Discussion Paper from the BJS Fellows Program.* September, Washington: U.S Department of Justice.

Manning, P. K. (2001). Theorizing policing: the drama and myth of crime control in the NYPD. *Theoretical Criminology, 5*(3), 315–344.

Maple, J. & Mitchell, C. (2000). *The crime fighter: How you can make your community crime free..* New York: Doubleday.

Martin, D. (2003). The politics of policing: managerialism, modernization and performance. In R. Matthews, & J. Young (Eds.), *New politics of crime and punishment.* Willan Publishing, Uffculme, Devon, pp. 227–249.

Marzullli J., & O'Shaughnessy, P. (2000, April 6). Cops feeling gun shy. *Daily News*, p. 17.

Maxfield, M. G., & Babbie, E. (2009). *Basics of research methods for criminal justice and criminology* (2nd ed.). Belmont, CA: Wadsworth/Cengage.

Mazerolle, L., Rombouts, S., & McBroom, J. (2007). The impact of Compstat on reported crime in Queensland. *Policing, 30*, 237–256.

McDonald, P. (2002). *Managing police operations.* Belmont, CA: Wadsworth.

McLaughlin, E. (2007). *The new policing.* London: Sage.

Messner, S. F., Galea, S., Tardiff, K. J., Tracy, M., Bucciarelli, A., Piper, T. M., et al. (2007). Policing, drugs, and the homicide decline in New York City in the 1990s. *Criminology, 45*(2), 385–413.

Moore, M. (2003). Sizing up Compstat: an important administrative innovation in policing. *Criminology and Public Policy, 2*(2), 469–494.

Moses, P. (2005, December 20). Something's missing: crime stat revelation. *Village Voice.* p. 1.

Neuman, W. L. (2000). *Social research methods: qualitative and quantitative approaches* (4th ed.). Boston: Allyn and Bacon.

New York City, Civilian Complaint Review Board. (2001). *Street stop encounter report: an analysis of CCRB complaints resulting from the New York Police Department's 'stop and frisk' practices.* Retrieved December 30, 2009 from: http://www.nyc.gov/html/ccrb/home.html

O'Connell, P. E. and Straub (1999). Why the Jails Didn't Explode, *City Journal*, 2, Spring, 28–37.

Parascandola, R. (2010, February 17). Cops urged to prove stats clean by citizens' crime commission. *New York Daily News*, p. 12.

Parascandola, R., & Levitt, L. (2004, March 22). Police statistics: numbers scrutinized. *New York Newsday*, p. 14.

Portas D., & Mason, J. (1996). *Collection and accuracy of police incident data.* London, David Keith Consultants.

Rashbaum, W. K. (2003, June 20). West Side

crime statistics were softened, police say. *New York Times*, p. B5.

Ritea, S. (2003a, October 24). Five NO [New Orleans] officers fired over altered crime stats. *The Times Picayune*, p. 1.

Ritea, S. (2003b, October 25). Crime, coercion and cover-up. *The Times Picayune*, p. 1.

Rosenfeld, R. (2007). Explaining the divergence between UCR and NCVS aggravated assault trends. In J. P. Lynch, & L. A. Addington (Eds.), *Understanding crime statistics*. New York: Cambridge University Press. pp. 53–69.

Rosenfeld, R., Fornango, R., & Baumer, E. (2005). Did ceasefire, Compstat and exile reduce homicide? *Criminology and Public Policy, 4*(3), 419–450.

Rosenfeld, R., Fornango, R., & Rengifo, A. F. (2007). The impact of order-maintenance policing on New York City homicide and robbery rates. *Criminology, 45*(2), 355–384.

Rothwell, G. R., & Baldwin, J. R. (2007). Whistle blowing and the code of silence in police agencies. *Crime and Delinquency, 53*(4), 605–632.

Ryan P. (2000) *NSW Police Service Future Directions, 2001-2005*, New South Wales Polices Service: Sydney, Australia

Safir, H. (1988). *The Compstat process*. New York: New York City Police Department.

Sanow, E. (2003). Compstat, the real cops on the dots. *Law and Order, 51*, 4–8.

Schick, W. (2004). Compstat in the Los Angeles Police Department. *Police Chief, 71*(1), 7–13.

Silverman, E. B. (2001). *NYPD battles crime: innovative strategies in policing* (revised ed.). Boston: Northeastern University Press.

Silverman, E. B., & O'Connell, P. (1998). The NYPD's new strategies. In A. Karmen (Ed.), *Crime and justice in New York City*. New York: McGraw Hill. pp. 61–72

Silverman, E. B., & O'Connell, P. (1999). Organizational change and decision making in the New York City Police Department. *International Journal of Public Administration, 22*(2), 217–250

Skogan, W., & Frydel, K. (Eds.). (2003). *Fairness and effectiveness in policing: the evidence.* (Committee to Review Research in Police Policy and Practices, Committee on Law and Justice, Division of Behavioral and Social Sciences and Education). Washington, DC: National Academies Press.

Solomon, F. F. (2003). *The impact of quality of life policing* (New York City Criminal Justice Agency Research Brief, No. 3) New York: NY: New York City Criminal Justice Agency.

Spitzer, E. (1999). *A report to the people of the State of New York from the Office of the Attorney-General.* New York: Attorney-General's Office.

Violent crime soars after police reporting failures. (2008, October 23). http://www.metro.co.uk

Walsh, W. F. (2001). Compstat: an analysis of an emerging police managerial paradigm. *Policing, 24*, 347–362.

Walsh, W. F., & Vito, G. F. (2004). The meaning of Compstat. *Journal of Contemporary Criminal Justice, 20*, 51–69.

Webber, R., & Robinson, G. (2003, July 7). Compstamania. *Gotham Gazette*, New York: Citizens Union. Retrieved from http://www.gothamgazette.com/article//2003 0707/997/432 on December 30, 2009.

Weisburd, D., Mastrofski, S. D., McNally A.M., and Greenspan, R. (2001) *Compstat and Organizational Change*, Report submitted to the National Institute of Justice by the Police Foundation.

Weisburd, D., Mastrofski, S., McNally, A. M., Greenspan, R., & Willis, J. J. (2003). Reforming to preserve: Compstat and strategic problem solving in American policing. *Criminology and Public Policy, 2*(3), 421–456.

Weisburd, D., Willis, J. F., & Greenspan, R. (2006). Changing everything so that everything can remain the same, Compstat and American policing. In D. W. Weisburd and A. Braga (Eds.), *Police innovation, contrasting perspectives*. Cambridge University Press.

Weissenstein, M. (2003, January 2). Call on NY's top cops: NYPD brass recruited by other cities to lower crime rates. *Newsday.* p. 18.

Whitehead, T. (2009, April 24). Top crime adviser admits public do not trust figures. *Daily Telegraph*, p. 4.

Willis, J. J., Mastrofski, S. D., Weisburd, D., & Greenspan, R. (2003a). *Compstat and organizational change in the Lowell Police Department*. Washington, DC: Police Foundation.

Willis, J. J., Mastrofski, S. D., & Weisburd, D. (2003b). *Compstat in practice: an in-depth analysis of three cities*. Washington, DC: Police Foundation.

Zink, R. (2004, summer). The trouble with Compstat. *PBA Magazine*. Retrieved from http://www.nycpba.org/publications/mag-04-summer/compstat.html on 6 May 2010.