Evaluation of an Early Intervention System at a Law Enforcement Agency

by
Robert Scott Russell

An Applied Dissertation Submitted to the
Abraham S. Fischler School of Education
in Partial Fulfillment of the Requirements
for the Degree of Doctor of Education

Nova Southeastern University
2014

ProQuest LLC.
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106 - 1346

Approval Page

This applied dissertation was submitted by Robert Scott Russell under the direction of the persons listed below. It was submitted to the Abraham S. Fischler School of Education and approved in partial fulfillment of the requirements for the degree of Doctor of Education at Nova Southeastern University.

Carolyn S. Buckenmaier, EdD Committee Chair	Date
David Weintraub, EdD Committee Member	Date
Mary Ann Lowe, SLPD Associate Dean	Date

Statement of Original Work

I declare the following:

I have read the Code of Student Conduct and Academic Responsibility as described in the *Student Handbook* of Nova Southeastern University. This applied dissertation represents my original work, except where I have acknowledged the ideas, words, or material of other authors.

Where another author's ideas have been presented in this applied dissertation, I have acknowledged the author's ideas by citing them in the required style.

Where another author's words have been presented in this applied dissertation, I have acknowledged the author's words by using appropriate quotation devices and citations in the required style.

I have obtained permission from the author or publisher—in accordance with the required guidelines—to include any copyrighted material (e.g., tables, figures, survey instruments, large portions of text) in this applied dissertation manuscript.

_____
Signature

Robert Scott Russell
Name

March 6, 2014
Date

Acknowledgments

The professional men and women of the study site law enforcement agency, from administrators to line members, each provided time, patience, and effort in assisting me with a virtual roadmap to the voluminous information sought for this study. I continue to be reminded how honored I am to serve beside them.

My chair and member, Dr. Carolyn Buckenmaier and Dr. David Weintraub, have provided unwavering support and guidance that enabled me to complete this final report. I also acknowledge my father and mother, James William Russell, Jr. and Betsy Ann Russell. I miss them both every day. I am still learning the lessons they offered me in life, and I know they would both be proud of this accomplishment. My thoughts and love also go to my loving wife, Lauren. As this chapter of my life closes, I live for the opportunity to experience so many more with her by my side.

Abstract

Evaluation of an Early Intervention System at a Law Enforcement Agency. Robert Scott Russell, 2014: Applied Dissertation, Nova Southeastern University, Abraham S. Fischler School of Education. ERIC Descriptors: Computer Software Evaluation, Crime Prevention, Law Enforcement, Police Community Relationship, Program Evaluation

The problem addressed through this program evaluation was that no formal study had been conducted regarding the implementation and effectiveness of the BlueTeam Program (BTP) within the law enforcement agency (LEA) serving as the study site. The BTP is a program that utilizes a computer application to track officer behaviors and alert administrators to potential trends in officer misconduct and complaints against officers. The program evaluation was guided by the process and product segments of Stufflebeam's (2003) content, input, process, and product model.

To conduct the evaluation, the researcher used a mixed methods approach for analyzing both qualitative and quantitative data. The perceptions of LEA stakeholders regarding the BTP, such as the sufficiency of staffing, budget, training, and ongoing support for effective implementation, were first collected. Quantitative data, consisting of archived, deidentified indicators of officer misconduct and complaints against officers acquired through the BTP, were then analyzed.

Findings of the study were that the BTP was effective in reducing incidents of officer misconduct and complaints against officers and for use in identifying which alerts were valid indicators of misconduct and complaints against officers. The one concern of stakeholders involving the BTP was limited nighttime vision; the recommendation for program improvement is that this shortcoming be addressed to determine possible solutions. Recommendations for future research involve the need for initial determinations, as well as formative evaluations, pertaining to the following three areas: (a) ascertaining the way in which the early intervention system will be used, (b) identifying the indicators of misconduct that will be tracked, and (c) determining the threshold at which the system will issue an alert.

Table of Contents

Chapter 1: Introduction

Members of law enforcement agencies (LEAs) experience challenges on multiple levels. Problems arising on an internal level, such as officer misconduct, result in both internal and external impact. Examples of officer misconduct are depicted by the media in various geographical regions. Officer misconduct became a national headline when 45 New Orleans police officers abandoned their posts during Hurricane Katrina (Associated Press, 2005). During the same storm, allegations were filed against five New Orleans officers for possible involvement in an unlawful shooting, leaving civilians for dead, and attempting to cover up the incident. The five officers were recently found guilty of civil rights and obstruction-of-justice violations for which they received sentences of between 38 and 65 years in federal prison (U.S. Department of Justice, Office of Public Affairs, 2012a). A similarly egregious pattern of misconduct was noted by members of the Federal Bureau of Investigation after the arrest of more than 100 allegedly corrupt police officers in Puerto Rico (Federal Bureau of Investigation, 2010). In a separate case, Lebovich and Rabin (2010) detailed the arrest of a Miami police officer who was involved in two fatal shootings.

Concerns of LEA administrators involving officer misconduct were not limited to the largest law enforcement departments. This was noted in a review of a press release issued from the U.S. Department of Justice Office of Public Affairs (2012b) which explained the demise of a West Memphis, Arkansas, officer convicted of criminal civil rights violations for the choking of a handcuffed arrestee. In review of another report, it was noted that the Florida Department of Law Enforcement (2009) released Year 2009 findings of 488 officers wherein action was taken in the form of suspension or revocation of officer certifications.

Officer misconduct results in complaints by citizens and the erosion of public confidence in the police. In recent years, early intervention systems had become a tool for addressing the problem of officer misconduct (Hassell & Archbold, 2010). In 2004, an LEA in a southern state implemented the BlueTeam Program (BTP), one form of early intervention system, as a means of addressing the problem of officer misconduct and complaints against officers. Through this program evaluation, the researcher examined the implementation and effectiveness of the BTP at the LEA within the research setting.

Statement of the Problem

The problem addressed through this program evaluation was that no formal study had been conducted regarding the implementation and effectiveness of the BTP within the LEA within the research setting. As a result, related information regarding officer misconduct and complaints against officers had not been collected and analyzed. The BTP was implemented to provide vital information as a preventive effort to reduce instances of officer misconduct and complaints against officers (LEA senior executive officer, personal communication, October 18, 2011). Without evidence of program effectiveness, LEA administrators were operating at a disadvantage in the area of strategic planning.

The research problem. A problem existed with officer misconduct at the LEA within the research setting. Incidents of serious official misconduct by officers were depicted in public records as well as the local and national media. Examples of official misconduct by police officers included the most egregious of incidents such as civil rights violations, bribery, and falsification of evidence (Gottschalk, 2011). Administrators of the LEA continued to address officer misconduct and the resulting negative consequences such as complaints against officers. As one approach for minimizing the

ongoing events involving officer misconduct, LEA administrators created the BTP in 2004. This evaluation study examined the implementation and effectiveness of the BTP. The BTP is explained in detail within a later section of this dissertation.

Serious incidents of officer misconduct continually occurred during the 9 years after the implementation of the BTP (LEA senior executive officer, personal communication, October 18, 2011). The most serious cases of officer misconduct resulted in officer dismissal from the LEA. Through this program evaluation, the researcher conducted the initial investigation into the implementation and effectiveness of the BTP as an early intervention system.

Officer misconduct, as used in this program evaluation, primarily involved selected behaviors such as the unnecessary use of physical force while fulfilling positions of public trust. The BTP, however, was additionally used to track other types of misconduct that can affect the relationship of police officers with fellow workers and citizens as well as the perceptions of the agency held by community residents. Related investigations can have detrimental effects on community relations (Dekmar, 2010). Investigations involving officer misconduct occur based upon complaints initiated by either the public or LEA administrators. Considering the potential impact, previous researchers found it valuable to include complaints against officers as a unit of analysis in the evaluation of programs to prevent officer misconduct (Hassell & Archbold, 2010).

Although the BTP was implemented to provide vital information as part of a solution to the problem of officer misconduct and complaints against officers, related concerns continued at the LEA within the research setting. Until this program evaluation was implemented, no formal study had been conducted regarding the BTP (LEA senior executive officer, personal communication, October 18, 2011). Without evidence of

program effectiveness, LEA administrators were operating at a disadvantage in terms of strategic planning. This evaluation examined the implementation and effectiveness of the BTP on the problem of officer misconduct and complaints against officers at the LEA.

Topic. The topic of this evaluation was the BTP program which was designed to identify officer misconduct in law enforcement. The evaluation was focused on the implementation and effectiveness of the BTP. The quantitative data collected for the evaluation were historical, public information acquired through the BTP as well as reports of officer misconduct and complaints against officers. In addition, qualitative data reflecting the perceptions of LEA stakeholders regarding the implementation of the BTP were collected and analyzed. The triangulation of data from multiple sources provided a foundation for comprehensive understanding and review of the program (Stufflebeam, 1999).

Research setting. The LEA is located in a southern state and holds jurisdiction for approximately 2,000 square miles. Based on a review of internal data, an average permanent population of approximately 200,000 residents, in addition to 100,000 visitors each year, is served through the LEA. The permanent population has been identified as a majority of non-Hispanic, White citizens, with the second and third most statistically significant race distributions being Hispanic and non-Hispanic Black. The most significant distribution of age categories within the population is between 35 and 54. The LEA operates under normal economic conditions with approximately 1,300 total employees including 918 certified officers. Within the jurisdiction of the LEA, the total annual arrests are approximately 15,000 for combined criminal violations. Of the 15,000 violations, homicide, robbery, and sexual assault crimes average approximately 5,000 per year.

The U.S. economic recession, according to representatives of the National Bureau of Economic Research, began in December 2007 (Temin, 2010). In the years following, during which the economic turndown continued, substantial impacts to the overall quality of life, similar to those in most communities across the nation, occurred within the jurisdiction of the LEA. Examples include foreclosure rates which increased from an average of 1,000 per year to more than 7,000 and unemployment rates peaking at almost 13%.

With each year of declining property values and subsequent reduction in tax base within the jurisdiction, LEA administrators were forced to operate with a cumulative budget reduction of nearly $25 million in tax dollars. More than $60 million of operating funds were lost when including the reduction in federal grants. This amount, which equates to approximately 40% of the overall prerecession budget, understandably caused LEA administrators to enact across-the-board cuts. In consideration that the largest line item in the LEA budget is employee salary and benefits, the progressive reductions resulted in unfilled vacancies of approximately 250 (27%) certified officers.

At the time this evaluation was conducted, approximately 220 officers were assigned to patrol duties; these officers responded to an average of 600,000 calls each year. Of those 600,000 calls, approximately 10% involved in-progress crimes or imminent threats to public safety and were designated as the utmost priority. Patrol officers issued approximately 8,000 traffic citations and 14,000 traffic warnings over the course of a year. Patrol officers encountered the most serious of criminal offenders within the jurisdiction, including approximately 200 registered sexual predators, 250 career criminals, and 1,400 documented members or associates of criminal street gangs. Through the daily risk exposure, approximately 40 patrol officers are assaulted each year.

Patrol officers are supervised by approximately 40 line supervisors whose chain of command includes approximately 30 midlevel managers and senior administrators.

The Program

Administrators of the LEA implemented the BTP in 2004 in an effort to reduce rates of officer misconduct and complaints against officers as well as to obtain evidence regarding the behaviors that are most predictive of misconduct and complaints. The decision to implement the BTP as a tool to reduce officer misconduct and complaints against officers was based, in part, on the review of literature by LEA administrators regarding the rising popularity of such systems (LEA senior executive officer, personal communication, October 18, 2011). Researchers such as Walker, Alpert, and Kenney (2001) had promoted early intervention systems for reducing instances of officer misconduct and related complaints. In addition, LEA administrators sought to meet new requirements of the Commission on Accreditation for Law Enforcement Agencies (2010) involving the implementation of early intervention systems. Related requirements were further established by the national accreditation commission of which the LEA maintained membership.

The software tracking tools utilized as part of the BTP are BlueTeam (Version 3.0) and IA Pro (Version 7.0); both were developed through CI Technologies (2010a, 2010b). These tools were chosen based on the belief of LEA administrators that they were the premier software platforms in the industry (LEA senior executive officer, personal communication, May 7, 2012). The specific research leading to this decision was conducted by an LEA program manager. No documents relating to the research were on file at the LEA.

The BTP is staffed by members of the internal affairs unit of the LEA and is a

portion of their assigned duties of addressing officer misconduct and complaints against officers. The BTP staff consists of two analysts and one program manager who are supervised by a program supervisor and senior executive officer. No budgetary adjustments were made for staffing, as the BTP was included within normal duties of selected staff within the internal affairs unit. Administrators of the LEA intended to implement the BTP with straightforward, core steps of (a) selecting BTP staff and ensuring proper knowledge, training, and time were provided for successful program implementation; (b) setting BTP data categories and thresholds based on staff feedback; (c) providing training to stakeholders potentially affected by the new program; (d) implementing the BTP as originally designed; and (e) facilitating ongoing feedback from BTP staff and stakeholders for needed program adjustment (LEA senior executive officer, personal communication, May 7, 2012).

The BTP tracks multiple categories of behavior and officer-involved incidents. The data are populated from real-time agency computer records. If data levels for an officer exceed a predetermined threshold in any category, an alert notifies members of the administration (see Appendix B). The categories of tracked data, as well as the threshold criteria, are defined by LEA administrators. When accessing the BTP computer portal, LEA administrators are able to view a snapshot of tracked officer data and corresponding levels of incident numbers leading to threshold maximums. The researcher created an illustration similar to the BlueTeam Dashboard as a visual reference for the reader (see Appendix A). Administrators of the LEA established seven requirements for effective management of the BTP:

1. Define which categories of data will be tracked by the BTP based on feedback from BTP staff.

2. Define and set a threshold of BTP alerts based on feedback from BTP staff.

3. Provide sufficient staffing for implementation and ongoing maintenance of the BTP.

4. Provide sufficient work-time allotments to BTP staff for effective implementation.

5. Provide sufficient training and familiarization to BTP program staff and LEA stakeholders.

6. Provide for ongoing feedback from BTP staff and stakeholders.

7. Evaluate and refine BTP as necessary.

Defining categories of tracked data. Administrators of the LEA worked collaboratively to identify the most appropriate officer behavioral and incident data for inclusion within the tracked categories of the BTP (see Appendix B). During the decision-making process, consideration was given to guidelines provided by Schultz (2011) for the International Association of Chiefs of Police. The guidelines outline 18 categories of potentially valuable data categories for inclusion within an early intervention system. The suggested categories include incidents such as (a) use of force; (b) discretionary arrests; and (c) complaints against officers from citizens, the agency, or coworkers (Schultz, 2011). Administrators of the LEA chose definable categories of tracked data to include 17 types of incidents such as (a) various complaints against officers, (b) the use of force including Taser and firearm deployment, (c) discretionary arrest charges, and (d) other data sets deemed valuable by LEA administrators (see Appendix B).

Defining alert thresholds. While attempting to decide the maximum number of incidents in each category that, if exceeded, would generate an alert to notify BTP staff,

LEA administrators discovered that this decision was completely discretionary to the individual agency (Walker, Milligan, & Berke, 2005). Based on input from BTP staff regarding agency norms, LEA administrators set various thresholds for the individually tracked data (see Appendix B), which, if exceeded within 12 months, would create a computer-generated alert to BTP staff (LEA senior executive officer, personal communication, May 7, 2012).

Staffing and work-time allotment. The staff members of the BTP were selected by LEA administrators based on their knowledge of each member's skill set, interest, and general proficiency with computer software applications. Administrators of the LEA made efforts to shift a portion of the normal workload of BTP staff members to other members of the internal affairs unit to allow for sufficient focus on the BTP.

Training and familiarization for BTP staff and stakeholders. Administrators of the LEA facilitated the training and orientation to the BTP staff; training was personally conducted by a representative from the BlueTeam (Version 3.0) and IA Pro (Version 7.0) software publisher, CI Technologies. The training consisted of three 8-hour sessions, provided over 3 days of time, and included the topics of software integration, functionality, and customization based upon requirements of LEA administrators. Subsequent to the training session, two meetings were held between LEA administrators and BTP staff members to address concerns, feedback, and ideas.

The BTP stakeholders at the LEA, supervisors and officers who would be affected by the new program, also required knowledge of its functionality. To address this need, stakeholders were provided BTP training and orientation in large groups. The training consisted of 4 hours of instruction facilitated by representatives of CI Technologies. Multiple 4-hour sessions were held to ensure all stakeholders were provided the

opportunity to attend. The content of the training included how to access the BTP, the goals of the new program, and the data and thresholds that had been designated by LEA administrators. The stakeholders were provided contact information for BTP staff members for use in addressing questions or concerns in a private forum.

Continued feedback from BTP staff and stakeholders. Members of the BTP staff are provided support from representatives of CI Technologies for technical assistance. As one form of support, BTP staff members attend an annual 3-day conference designed to update agencies utilizing the program software. Administrators of the LEA allow for communication of additional concerns or issues through an open-door policy that is not required to follow the normal chain of command. This policy is in place to assist stakeholders in addressing concerns or issues with supervisors and officers of their choice; discussions are not shared with BTP staff. The open-door policy is additionally used in other cases to address issues that an agency member feels were not resolved satisfactorily.

Evaluate and refine BTP as necessary. As noted in the literature, the effective implementation of the BTP could only be attained through frequent evaluation and adjustments to data and thresholds collected by the software tools (Bertoia, 2008; Walker et al., 2005). However, budgetary crises had prevented such an evaluation from occurring within the LEA. A lack of funding to employ an outside evaluator, the desire to know the effectiveness of the BTP, and questions involving whether the data and thresholds should be adjusted led to the willingness of LEA administrators to agree to the current evaluation (LEA senior executive officer, personal communication, May 7, 2012). By conducting the evaluation, the researcher was assisting the agency while also completing necessary requirements for a doctoral dissertation.

Justification for the Program Evaluation

Walker (2005) reported that ongoing problems of officer misconduct and complaints against officers have prompted LEA administrators across the United States to design solutions to reduce these issues. Although LEA administrators within the research setting had attempted such reform, events of officer misconduct and complaints against officers had continued (LEA chief executive officer, personal communication, July 21, 2011). Such events have broad impact; as Prenzler (2009) described, "Police misconduct has a deeply corrosive effect on society, undermining the system of demographic authority and threatening the security of ordinary citizens" (p. 1).

While continuing to address officer misconduct and complaints against officers, an equally concerning factor for the LEA administrators was the void of knowledge regarding the effectiveness of the BTP (LEA chief executive officer, personal communication, July 21, 2011). A key element of a successful early intervention system is ongoing evaluation (Bertoia, 2008). As Walker et al. (2005) affirmed, "It is necessary to continually evaluate the choice of data being collected, how useful they are, whether new data should be collected, and whether the data being collected are of high quality" (p. 34).

The researcher's background involves practical and academic experience in both criminal justice and research. While an undergraduate student, the researcher was employed by the Florida Department of Law Enforcement. Among the job responsibilities was the drafting of the initial version of a new Florida law to adopt the National Child Protection Act (U.S. Government Accountability Office, 1997). The researcher's first participation in a published research study involved the topic of police use of force. The study involved both qualitative and quantitative research methods,

including survey research of the population of Florida municipal LEAs regarding the use of force and oleoresin capsicum deployment (Florida State University, 2000). The research was supported through a grant from the National Institute of Justice and facilitated by faculty of the Florida State University.

Within qualitative research methods, the use of focus groups and surveys has been found to provide valuable in-person and written interactions with researchers (Creswell, 2012). One valuable opportunity during focus group interviews is the ability of the researcher to observe nonverbal cues and communication exhibited by participants, whereas written responses can be examined for consistency and truthfulness (Eriksson & Kovalainen, 2008). The researcher has been an active, certified law enforcement officer within the LEA within the research setting for 13 years. During the tenure at the LEA, the researcher authored the internal policy defining procedures to be utilizing during eyewitness and photographic identifications. While serving in an assignment as a detective, advanced certifications were obtained in the areas of kinesics, investigative discourse analysis, and deception detection during both personal interviews and the examination of written statements.

Deficiencies in the Evidence

Prior to conducting this program evaluation, the BTP had not received close scrutiny to assess its effectiveness (LEA senior executive, personal communication, October 18, 2011). Although early interventions systems such as the BTP had become more prevalent, few researchers have studied the effectiveness of the systems (Lersch, Bazley & Mieczkowski, 2006). While LEA administrators continued to address officer misconduct and the related issue of complaints against officers, the lack of knowledge regarding the BTP effectiveness was significant. From the classical work of Johnson

circa 1759 (as cited in Gaffigan & McDonald, 1997), “Integrity without knowledge is weak and useless, and knowledge without integrity is dangerous and dreadful” (p. 86).

Intended Audience

The stakeholders in this study are defined as the administrators and officers of the LEA as well as the citizens within the community. Under normal economic conditions, the LEA is operated with approximately 1,300 employees among which are 918 certified officers. Under the current economic recession, approximately 220 officers are assigned to patrol duties. These officers respond to an average of 600,000 calls annually. Patrol officers are supervised by approximately 40 line supervisors whose chain of command includes approximately 30 midlevel managers and senior administrators. Administrators of the LEA requested this evaluation be conducted to discover “to what level the BTP is effective in reducing officer misconduct and complaints, and specifically which data components are most valuable as indicators of officer misconduct and complaints” (LEA chief executive officer, personal communication, July 21, 2011). Such answers could allow administrators at the LEA, as well as those employed through other agencies, to target outcome indicators based on statistically valid data (Gibbs & Kendrick, 2011).

Definition of Terms

Alert. An alert is a computer-generated notification to LEA administrators that occurs when an officer has generated incident data within the BTP that exceeds predetermined thresholds.

BlueTeam Program (BTP). The early intervention system used at the LEA within the research setting is the BTP. The BTP is a program that utilizes software tools to track behaviors defined by administrators and generates notifications when predetermined thresholds are exceeded. The BTP is encompassed within policies and

procedures designed to correct identified misbehaviors; these procedures include officer retraining, counseling, probation, or disciplinary action when deemed necessary by the administration.

BTP administrators. Members of the LEA who have been assigned to tasks involving the implementation and operation of the BTP are referred to as BTP administrators.

BTP line supervisors. BTP line supervisors are members of the LEA who hold supervisory rank and have been assigned to patrol duties. The duties of such members involve the implementation and operation of the BTP from an end-user perspective.

Complaint. The formal filing by a citizen, an LEA officer, or an LEA administrator reporting alleged actions or behaviors by an LEA officer is a complaint.

Early intervention system. This system is a "data-based management tool designed to identify officers whose performance exhibits problems, and then to provide interventions, usually counseling or training, to correct those performance problems" (Walker, 2005, p. 3). The early intervention system was originally termed an *early-warning system,* and the terms are used interchangeably within current literature.

Law enforcement agency (LEA). An LEA is defined as a "legally constituted governmental entity having mandated responsibilities to enforce laws and having personnel with general or special law enforcement powers" (Commission on Accreditation for Law Enforcement Agencies, 2010, para. 3).

LEA administrators. This term refers to members of the LEA who hold rank at the executive level and are empowered to make decisions regarding policy, operations, and strategic planning for the agency.

LEA officer. An individual who is a member of the LEA in any capacity, paid or

unpaid, who is directed to behave within law and policy is termed an LEA officer.

LEA certified officer. This term refers to an individual who is employed within the LEA and, while employed, has the powers of arrest and related duties of a law enforcement officer.

Member checks. This process involves clarifying scripted interview responses with participants for the purpose of ensuring the accuracy of data (Mills, 2010). In this program evaluation, member checks were conducted to ensure handwritten responses to the focus group discussion were accurately documented.

Officer misconduct. Officer misconduct, as used in this program evaluation, was defined by Kane and White (2009) as the unnecessary use of physical force while fulfilling positions of public trust. Misconduct, which can additionally include illegal, immoral, or improper conduct occurring while police officers are off duty, is believed to affect the relationship of police officers with fellow workers and citizens as well as the perceptions of the agency held by community residents (Braga, 2010; Kane & White, 2009; Spinelli, 2010).

Triangulation. The triangulation of data involves validating data sources by incorporating a variety of data collection strategies (McMillan & Schumacher, 2009).

Trustworthiness. The trustworthiness of qualitative data involves the degree to which researcher bias, subjectivity, and values affect the findings acquired through the study (Chenail, 2011; Mills, 2010).

Workload. The workload consists of documented activity by LEA officers that involves citizen interaction or duty-related tasks. The inclusion of workload during data analysis of misconduct-related studies secures higher rates of validity in results (Harris, 2010b).

Purpose of the Evaluation

The purpose of this program evaluation was to determine the effectiveness of the BTP implementation and operation regarding the reduction of officer misconduct and complaints against officers at the LEA within the research setting. At the onset of the evaluation, the information gained from the investigation was expected to be useful as research-based results for reference by LEA administrators as guidance when deciding whether to continue or adjust the current parameters of the BTP. The acquired information was also expected to be useful in assisting in the decision to continue or adjust current strategies of counseling, training, and discipline employed in concert with the BTP to address the problem of officer misconduct, complaints against officers, and relative officer performance. The goal held by LEA administrators, both currently and during the planning and implementation of the BTP, has been to reduce officer misconduct and complaints against officers (LEA chief executive officer, personal communication, July 21, 2011).

Rationale for a Program Evaluation

The role of evaluation in this study was to assist in narrowing the gap of knowledge regarding the effectiveness of a program relative to the reduction of officer misconduct and complaints against officers and to assist in providing valuable information for future decisions on how to adjust a program for enhanced operation (Froggatt & Hockley, 2011). Within government organizations, such evaluations can be a vital element in strategic planning (U.S. Government Accountability Office, 2012). When conducting an evaluation within the law enforcement community, where integrity is held at an upmost position, results of program evaluations should provide unbiased, straightforward information while also utilizing public funds responsibly (Shaw, Greene,

& Mark, 2007).

Summary

Administrators of the LEA within the research setting had experienced events of officer misconduct and complaints against officers as also noted within many agencies across various boundaries and jurisdictions (Treverton, Wollman, Wilke, & Lai, 2011). When attempting to manage the impact of officer misconduct and related complaints, LEA administrators implemented the BTP as an early intervention system. In an effort to determine whether the BTP was effectively implemented and was producing the desired outcomes, LEA administrators were interested in obtaining data to determine the efficacy of the BTP. This evaluation was conducted to serve as a basis for future decisions regarding this program.

Chapter 2: Literature Review

Officer misconduct and complaints against officers have continued to cause sufferable challenges for LEA administrators for decades. This chapter begins by exploring the literature regarding the history of policing. The root causes of police misconduct are reviewed; this discussion includes the tools and previous strategies used to identify and reduce instances of misconduct. Reviewed are findings of how early intervention systems are designed, the data they should track, and the conditions under which they are most effective. In an effort to determine the most appropriate method to evaluate an early intervention system, the researcher reviewed literature on evaluation models. As a result of the review, the researcher determined that the process and product segments of the context, input, process, and product (CIPP) model of evaluation (Stufflebeam, 2003) would best assist in collecting the information needed for this study.

History of Policing in America

Officer authority and citizen expectations. Police officers are expected to protect citizens and their property and, most importantly, uphold the civil and political rights of citizens (Archbold, 2005). Embodying the values written in 1787 within the Preamble to the U.S. Constitution, police officers are empowered to "establish justice, insure domestic tranquility, provide for the common defense [and] promote the general welfare" of the citizens and communities served (U.S. Senate, 1994, Preamble section, para. 1). Citizens grant the authority given to police officers and expect officers to demonstrate their virtue by holding such authority without abusing it (Macaulay, 1908). While depicting the challenges in American policing, Goldstein (1977) explained that, due to the role police officers play in a free society, their existence is an anomaly. In essence, although citizens want a free society, they have empowered the police to restrict

and enforce the limits of such freedom (Goldstein, 1977).

Historical eras and styles of policing. Treverton et al. (2011) wrote of the historical benchmarks of police organizations as four eras progressing to the current date: (a) political, (b) reform, (c) community, and (d) intelligence based. The political era contained voluminous amounts of corruption that provoked the onset of the reform era through citizen demands. The community era, focusing on relationships between officers and the citizens, is still apparent in many areas yet has become enhanced by the advancements in technology available to police agencies. With technology, the movement toward intelligence-based policing includes the utilization of computer hardware and software to not only track crime trends but to also track officer behavior in an effort to identify and correct misconduct (Klockars, Ivković, & Haberfeld, 2004).

Police Misconduct

From the onset of policing and the creation of organized police departments, concerns of officer misconduct and abuse of power have occupied the thoughts of community members and police administrators (Frank, 2009). Hughes and Andre (2007) stated the importance of successfully identifying problematic behaviors early in an officer's career to prevent serious incidents that could be detrimental to the public image of an LEA and continued success in the mission to serve the community. Furthermore, Hughes and Andre noted that a general consensus among the community of police researchers is that approximately 90% of misconduct occurrences within law enforcement can be attributed to a mere 10% of the officer population. The history of policing also suggests that misconduct continues despite many decades of effort and the use of numerous intervention strategies (Walker & Macdonald, 2009).

The literature concerning the topic of police misconduct has emphasized the

salient role of integrity. Gottschalk (2011), for example, stated, “Integrity in public office demands open and transparent decision making and clarity about the primacy of a public official’s duty to serve the public interest above all else” (p. 171). Huberts, Kaptein, and Lasthuizen (2007) similarly reported that integrity is vital to the functioning of public entities in order to (a) strengthen the confidence of stakeholders; (b) reduce the need for external investigations, regulations, and conflict; and (c) elicit the cooperation of stakeholders. Pogarsky and Piquero (2004) similarly reported that integrity is the key fundamental requirement for police officers to fulfill a variety of job-related activities requiring the confidentiality and trust of stakeholders. Rossler and Terrill (2012) further underscored the role of integrity by stating that public legitimacy suffers with the lack of officer integrity. In a very pragmatic definition, Grogan (2011) described integrity as the actions people take when they believe no one can see them.

The continuum between the integrity and corruption of police officers is inclusive of numerous points that are critical to both individuals and entities. The contrast, when moving from the presence of integrity toward its absence and culminating at corruption, is a disparity that can bring financial loss and irreparable damage to the reputations of individual officers and police departments (Huberts et al., 2007). Vigneswaran (2011) further stated that corruption undermines the ability of police officers to enforce the law and weakens the deterrence of crime. These collective repercussions support the claim of Gottschalk (2011) that officer misconduct is a matter of grave concern within communities across America.

Police misconduct falls on a continuum ranging from minor violations of policy and practice to departmental corruption. McElvain and Kposowa (2004) reported that charges involving violations can stem from (a) citizen complaints, (b) supervisory

observations or questions, or (c) internal affairs investigations. Regardless of severity along the continuum, related activities are vulnerable to personal biases and external influences because of the limited supervision that officers receive while fulfilling job duties (Hickman & Piquero, 2009; Pogarsky & Piquero, 2004).

Vigneswaran (2011) explained that a significant phenomenon of concern in the field of criminology and policing studies is that of police misconduct. Nickels (2007) defined police misconduct as infractions involving either unethical actions or the lack of discretionary action. Jiao (2010) asserted that officers do not always become involved in misconduct for the purpose of acquiring financial or material gain. In contrast, however, Jiao defined corruption as misbehaviors in which officers engage explicitly for the purpose of acquiring financial or material gains. Huberts et al. (2007) more succinctly defined corruption through two viewpoints; the first is a micro view, depicting corruption as a behavior that deviates from the innate duties of the public role in favor of gaining benefits. A macro vision of corruption, conversely, describes corruption as a violation of the moral norms and values of the community as a whole (Huberts et al., 2007). As also noted by Gottschalk (2011), police corruption demonstrates a lack of integrity.

A veil of secrecy guards occurrences of police misconduct and corruption (Stinson, Liederbach, & Freiburger, 2010). No comprehensive data are available regarding such crimes, and no data on criminal arrests of law enforcement officers are maintained by federal officials (Stinson et al., 2010). Harris (2010a), as well as Lentz and Chaires (2007), similarly described the dearth of literature regarding police misconduct and related discipline. Along the same lines, Richards (2010) stated, “Police organizations form ethics that have a distinctive and often unique character” (p. 223). Richards additionally described law enforcement as a “context of evasive particularity . . .

of practice" (p. 223). Alpert and Noble (2009) simply described the unique secrecy and related ethics contributing to the difficulty of identifying acts of misconduct and corruption as "organizational tolerances" (p. 239).

A review of the literature revealed numerous examples of documented officer misconduct. Based on common personnel practices and legal guidelines, misconduct can be divided into the two categories of official and personal, with official misconduct being the more severe of the two and having the more significant ramifications on the departmental reputation (Gottschalk, 2011). Each category is discussed in the following text.

Official misconduct. Johnson and Bridgmon (2009) reported that the misconduct of police officers has "a profound and negative impact on the citizenry's faith in the criminal justice system which can lead to a myriad of issues related to police-community relations" (p. 197). Hassell and Archbold (2010) identified two primary categories of official misconduct, and several examples of each were noted in the review of the literature. The first involves violations of police policy and practice; the second category is the dereliction of duties. Examples within each of these categories are provided in the following text.

Violations of police policy and practice noted in the review of the literature include a total of 32 practices. Frequently cited examples were the unnecessary use of force (Gottschalk, 2011; McClellan & Gustafson, 2012; McElvain & Kposowa, 2004) and assault (Gottschalk, 2011; Richards, 2010). Other common examples were fraud and theft (Gottschalk, 2011; Huberts et al., 2007; Richards, 2010); acceptance of gratuities and favors (Gottschalk, 2011; Huberts et al., 2007); and conflicts of interest which may involve assets, jobs, or gifts (Gottschalk, 2011; Huberts et al., 2007). Gottschalk (2011)

also identified extortion, coercion of confessions, fabrication of evidence, commitment of perjury through false testimony, bribery, nepotism, violating the rights of others, and the waste and abuse of resources. Hassell and Archbold (2010) additionally identified harassment, verbal threats, the improper use of actions during arrests, and discourtesy to citizens. Huberts et al. (2007) cited violence against suspects or citizens. Stinson et al. (2010) additionally identified petit and grand larceny, drug trafficking, extortion, bribery, gambling, narcotics, insurance fraud, burglary, receipt of stolen property, intimidation, perjury, and violations of civil rights and weapons laws. Hassell and Archbold (2010) further identified discourtesy to citizens as official police misconduct.

Dereliction of duties varies, based upon local policies; furthermore, the review of the literature indicated fewer examples in this category than in the category of violations of policy and practice. Dereliction of duties includes the failure to (a) properly document evidence (Gottschalk, 2011; Hassell & Archbold, 2010), (b) present identification to citizens (Hassell & Archbold, 2010), (c) use appropriate investigative methods (Huberts et al., 2007), (d) patrol assigned area (Stinson et al., 2010), and (e) take police action against observed crime (Stinson et al., 2010). Additional examples include the failure to book evidence properly, complete written reports when required, and maintain appropriate uniform and appearance (Gottschalk, 2011). Dereliction of duties also includes obstructing justice (Stinson et al., 2010).

Personal misconduct. Miller (2010) depicted a law enforcement officer as someone who accepts both professional and moral obligations "different from and additional to the moral obligations these individuals had before entering the profession" (p. 242). Discussion in the literature involving personal misconduct similarly resonated of integrity, values, and culture. As another example, Pogarsky and Piquero (2004) stated

that misconduct is rooted within community norms and values. Furthermore, Pogarsky and Piquero reported that behaviors of police officers are typically influenced by both the formal and informal cultures of the agency.

Gottschalk (2011) described personal misconduct as "private time misconduct" (p. 171); Huberts et al. (2007) similarly described personal misconduct as inappropriate or immoral behaviors occurring while off duty. Stinson et al. (2010) reported that slightly more than half the misconduct of police officers occurs while off duty and falls within the category of personal misconduct. Examples of personal misconduct include driving while intoxicated, domestic assault, predatory sex offenses, drug abuse, both simple and aggravated assault, and welfare fraud (Stinson et al., 2010). Tyler (2010) additionally identified the failure to heed traffic laws, involvement in illegal immigration, the nonpayment of taxes, and gambling as examples of personal misconduct.

Characteristics of high-risk officers. Harris (2010b) stated, "Criminologists have all but abandoned the notion of prospective identification of career criminals" (p. 224). This belief is founded in the fact that officers who make more arrests and issue more citations, when compared to their peers with limited public exposure, receive a higher number of complaints (Hassell & Archbold, 2010; McElvain & Kposowa, 2004; Miller & Davis, 2008). Furthermore, police officers encounter the public frequently, but not equally, across society as reflected by neighborhood needs; officers with high exposure to low-socioeconomic and high-minority neighborhoods receive the majority of complaints (Hassell & Archbold, 2010; McElvain & Kposowa, 2004; Miller & Davis, 2008). In spite of these beliefs, numerous researchers have identified several characteristics that may exemplify the characteristics placing police officers at risk of misconduct or corruption. These collective characteristics can be used as a profile of

high-risk officers.

Hassell and Archbold (2010) reported that race is not a determining factor in misconduct, as results are mixed involving White officers and officers of color. Hassell and Archbold, as well as Pogarsky and Piquero (2004), however, emphasized the belief that males are more prone to commit misconduct than their female peers. Furthermore, a strong association exists between prior and future offenses, underscoring the likelihood of repeat offenses among first-time offenders (Hassell & Archbold, 2010). Mixed opinions were reported involving the effects of years of service. Pogarsky and Piquero reported that officers with 3 or more years of experience are more likely than their novice counterparts to be involved in "minor" transgressions (p. 376). Hassell and Archbold, however, further stated that younger officers are more likely to have complaints filed against them.

The number of complaints filed against police officers is reduced commensurate with the years of formal education officers have acquired (Hassell & Archbold, 2010). Moreover, the larger number of complaints is generally filed against officers who (a) are more aggressive, (b) have issued the larger number of citations and arrests, and (c) have interrogated more suspects (Hassell & Archbold, 2010). Pogarsky and Piquero (2004) reported the existence of a strong association between impulsivity and officer misconduct; furthermore, a subgroup of officers appears to hold no fear of sanctions and has a high propensity toward impulsivity. In conclusion, the personal characteristics determining the profile of a high-risk officer appear to be consistent with an impulsive, aggressive male with prior offenses and a lower level of education.

Tools for identifying officers at potential risk. Practices involving officer screening can be traced back to London in 1829 (Cuttler & Muchinsky, 2006; Forero,

Gallardo-Pujol, Maydeu-Olivares, & Andrés-Pueyo, 2009). Although mental ability and selected aptitudes have been assessed in America for over 75 years, it was not until 1960 that the personality factors of candidates were initially assessed (Cuttler & Muchinsky, 2006). In light of reductions in public coffers across the country, Wilson and Heinonen (2011) emphasized the critical need for relevant, evidence-based staffing practices supported by data analysis.

The purpose of identifying effective identification tools is twofold. First, the tools are useful in detecting existing corruption; second, identification tools may be instrumental in preventing officers from initially engaging in corruption (Prenzler, 2006). Vigneswaran (2011) further reported that effective screening devices have been developed to (a) detect applicants who may be prone to corruption, (b) define the effects of social contexts and sociopolitical transformations, (c) track changes in the attitudes of officers toward corruption, and (d) measure the efficacy of management and oversight systems.

In addition to measures of cognitive performance and aptitude, psychological screening with a standardized assessment is a common practice (Caillouet, Boccaccini, Varela, Davis, & Rostow, 2010; Cuttler & Muchinsky, 2006). The Minnesota Multiphasic Personality Inventory–2 is one of the prevalent assessments used in the United States (Cuttler & Muchinsky, 2006). Subtests include the evaluation of four factors: (a) aggression, dominance, and a tendency to intimidate others; (b) risk taking, impulsivity, and boredom with routine; (c) psychotic symptoms, anxiety, depression, and symptoms of paranoia, schizotypal, and borderline personality disorders; and (d) worry, self-criticism, and guilt (Cuttler & Muchinsky, 2006).

A second approach frequently used for measuring influences of personality

involves the five-factor model (Cuttler & Muchinsky, 2006; Forero et al., 2009; Ono, Sachau, Deal, Englert, & Taylor, 2011). The five-factor personality taxonomy, in order of importance in screening for police officers, measures (a) conscientiousness, (b) neuroticism, (c) extraversion, (d) agreeableness, and (e) openness to experience (Cuttler & Muchinsky, 2006; Ono et al., 2011). The subtest of conscientiousness examines traits of organization, responsibility, carefulness, and work ethic of candidates (Barrick & Mount, 2012). The subtest of neuroticism examines the behavioral traits of candidates involving insecurity, anger, worry, and insecurity. The subtest of extraversion examines the traits of being social, agreeable, and interactive with others. The subtest of agreeableness examines the tendency of candidates to be adaptive, understanding, and tolerant of others. The subtest of openness to experience examines the tendency of candidates to be open minded, culturally diverse, and without prejudice of others (Barrick & Mount, 2012).

A third approach for predicting the ethical behaviors of candidates has been developed; this approach involves emphasizing the importance of the four constructs of hope, optimism, efficacy, and resilience (Walumbwa, Peterson, Avolio, & Hartnell, 2010). Efficacy refers to exuding the confidence to expend the necessary effort to succeed; hope reflects the ability to persevere toward a goal (Walumbwa et al., 2010). Optimism involves a positive expectation involving success, whereas resilience reflects the ability to sustain and bounce back when faced with problems and adversity (Walumbwa et al., 2010).

Once agency leaders identify valid approaches for measuring personality and predicting behaviors of officer candidates, the use of early intervention or early warning systems was recommended as the most promising approach for identifying officers with a

propensity for misconduct (Hassell & Archbold, 2010; Pogarsky & Piquero, 2004). Hassell and Archbold (2010), however, urged supervisors to make judgments by comparing activities of officers assigned to similar geographic areas and shifts to ensure that citizen complaints are valid. One drawback, as noted by Schwartz (2010), is that the installation and initial functioning of early intervention systems can take years to be effectively established.

Police accountability. The management of behaviors of police officers is "an important but frequently overlooked function of police organizations" (Wilson & Heinonen, 2011, p. 280). Accountability, however, is threatened by the embedded silence within the culture of police work that is historically passed from veteran officers to new recruits (Gottschalk, 2011; McElvain & Kposowa, 2004; Shane, 2010). The code of silence strongly protects officers from repercussions involving the excessive use of force and acceptance of gratuities and favors (Gottschalk, 2011).

It is also difficult to generate reliable data for use in addressing the illegal and socially stigmatized behaviors defining corruption in policing (Vigneswaran, 2011). In studies of police misconduct and crime, Gottschalk (2011) and Stinson et al. (2010) underscored the need for effective and ethical leadership yet indicated that the management of law enforcement officers historically has been challenged by the culture of officer solidarity and silence. The protection of members within the law enforcement community, ensured by the unwritten code of silence, minimizes the likelihood of misconduct surfacing and often protects officers from formal discipline (Gottschalk, 2011).

Richards (2010) similarly referenced the embedded culture of law enforcement as "loyalty" (p. 221), which is "a thick . . . concept" (p. 224). To depict the critical role of

loyalty, Richards likened the quality to "courage, integrity, and pride" (p. 221). Miller (2010) further described the organizational culture of police officers as a "pervasive ethos or spirit" (p. 243) of loyalty. In spite of related challenges, however, supervisors must establish effective accountability systems (Ross & Bodapati, 2006; Wilson & Heinonen, 2011).

Three approaches for establishing accountability systems surfaced in the literature review. First, Archbold (2005), Hassell and Archbold (2010), as well as Simmons (2008), recommended that supervisors maintain a record of formal and informal citizen complaints and develop external accountability measures such as citizen oversight boards. Simmons further accentuated the need for police to be included on the oversight boards. Second, emphasizing the risk exposure to police administrators and entities, Ross and Bodapati (2006) underscored the need for both ongoing training and risk-management strategies in efforts to reduce related liability. Third, similar to the traditional problem-solving model, Archbold identified five steps to risk management: (a) identifying "risks, frequency of exposure, and severity of losses"; (b) exploring "alternative methods for handling exposure to risk"; (c) selecting "appropriate methods for managing exposure to risk"; (d) implementing identified changes in procedures; and (e) evaluating the selected methods and making adjustments in an ongoing process (p. 31).

Impacts of Officer Misconduct

Complaints against officers. As noted in a roundtable discussion by law enforcement leaders throughout the United States, many LEA administrators continue to strive to reduce complaints against officers related to incidents of officer misconduct (U.S. Department of Justice, 2011). High numbers of complaints can impact public

confidence in the police (Gaffigan & McDonald, 1997), yet the value assigned to the number of complaints has been debated in the literature. Historical studies determined that approximately 30% of citizens who felt mistreated by the police filed reports (Harris, 2010a). In 2010, researchers described the trend that a larger percentage of citizens is now likely to come forward yet questioned the measurement of complaints without allowing for workload (Hassell & Archbold, 2010). Researchers have asserted there is a risk from incomplete assessment, as the complaints may result from an excessive workload and not true officer misconduct (Hassell & Archbold, 2010).

Regardless of whether a complaint of officer misconduct is confirmed, the impact can be detrimental to public confidence in the agency and the effectiveness of its policing efforts (Dekmar, 2010). Research findings by Johnson and Bridgmon (2009) included the discovery that the majority of the most serious complaints leading to federal indictments result from acts of local police officers. Hassell and Archbold (2010) reported that including complaints as a unit of analysis in the evaluation of programs of officer-misconduct prevention is imperative in securing accurate measurement.

Public trust. Prenzler (2009) pointed out the deleterious effect of police misconduct on society and went on to say that ordinary citizens have had their security threatened by such misconduct. Representatives of the U.S. Department of Justice, Office of Community Oriented Policing Services (2009) expressed that misconduct undermines the democratic authority of police officer:

> The building and maintenance of trust takes a great of continuous effort. Unfortunately, the ethical work of thousands of local law enforcement officers is easily undone by the actions of one unethical officer. In short, the integrity of the police will always dictate the level of public trust. (p. 3)

Regarding the importance of public trust and the future success of policing,

representatives of the International Association of Chiefs of Police (2004) explained that the public perception of police conduct has an impact on the willingness of citizens to cooperate with officers and obey the law. Enhancing the importance and challenge for police agencies to build community trust is the factor that often the communities with the most distrust in police officers are those who suffer high crime rates and are most in need effective policing (U.S. Commission on Civil Rights, 2000). At the epicenter of integral constructs needed for successful policing is the element of public trust (Fitch, 2011).

Liability. Administrators of LEAs often track several categories of behavior, including officer workload, in an effort to accurately identify the most injurious behaviors of officer misconduct. The early identification of these behaviors could avert legal action against an agency (Bertoia, 2008). This process was detailed by Kinnaird (2008) who explained, "In almost all civil litigation cases against the police, the organization must prove that it has used its control-documentation policies to anticipate the risk of negligent opportunities by officers" (p. 34). Similarly, agencies risk litigation if internal policies are not effective in identifying behaviors leading to officer misconduct (Walker & Macdonald, 2009). A roundtable discussion of law enforcement leaders revealed a strong desire for outcome evaluations of related programs to better develop evidence-based policies that prevent litigation (U.S. Department of Justice, 2011).

Economic consequences. Beyond the potential for costly litigation caused by officer misconduct, LEAs must absorb the costs related to the termination of an officer due to egregious acts of misbehavior. Orrick (2008) stated that a generalized estimate of losing an employee can range from 1 to 5 times the total salary of the employee. Because of notable deficiencies in officer positions under ongoing budget restraints, the retaining of every officer is crucial (Police Executive Research Forum, 2010).

Previous intervention efforts. Many attempts have occurred to create a more reliable form of accountability for officers and to implement preventative measures to limit acts of officer misconduct. Efforts have included (a) creating strict military structures, (b) providing enhanced officer training, (c) deterring behaviors with punitive discipline, and (d) attempting to professionalize the career of policing (Walker, 2010). The most salient strategy for police reform has occurred through various efforts to professionalize the image and identity of policing as a practice (Treverton et al., 2011). With this reform were efforts to develop organizational structures such as chain of command, written policies, and supervisor span-of-control designations (Treverton et al., 2011). In the wake of many publicized events of officer misconduct, several areas of failure are identifiable within the movement to professionalize policing (Frank, 2009).

One specific failure identified in years of police gravitation toward professionalism is the neglect to implement or maintain effective personnel-evaluation systems (Walker, 2005). One example of this type of failure is the 1999 Rampart event within the Los Angeles Police Department which was held as one of the most professional agencies in the country prior to the event (Walker, 2005). The scandal involved confirmed acts of corruption and misconduct by more than 20 officers and resulted in over 100 convictions stemming from arrests of the officers being overturned in court (Kaplan, 2009). The inquiry board of the department further found that personnel evaluations at various levels within the organization lacked credibility (Walker, 2005).

Early Intervention Systems

The use of early intervention systems is a method noted in the literature regarding efforts to reduce officer misconduct. As stated by Walker et al. (2001), “Perhaps the most increasingly used tool for quelling misconduct is the early-intervention or early-warning

system" (p. 13). As Noble and Alpert (2009) reported, "Early warning systems were endorsed by the U.S. Civil Rights Commission in 1981 and by the Justice Department conference on police integrity in 1996" (p. 286). These approaches are designed to identify the need for interventions, such as counseling or supplemental training, on behalf of officers who are frequently the subject of complaints.

The approach from a private company is largely similar to the law enforcement field in hopes of correcting employee behavior for alignment with the mission and goals of the organization (Mathis & Jackson, 2007). Stressing the need for LEA administrators to implement such systems, McDonough (2011) explained the crucial role of an early intervention system to assure the wellbeing of officers throughout their years of service. Walker et al. (2001) estimated that half of all LEAs in areas with populations greater than 50,000 citizens are either utilizing or preparing to implement a system. Bertoia (2008) asserted that early intervention systems can be effective when used properly and placed as part of an overall strategy to reduce police misconduct yet cautioned LEA administrators that success requires ongoing attention. Walker (2005) addressed this caution by stating that such systems are effective and have great potential when properly used yet are extremely complex and, without proper maintenance, can have a negative impact.

Software applications for early intervention systems. As technology in law enforcement continues to be developed, computerized programs that assist in the detection of officer misconduct have become more prevalent (Walker, 2005). For example, CI Technologies (2010a, 2010b) offers LEA administrators two software tools that integrate with each other as well as with supplemental data collected throughout the agency. The applications, titled BlueTeam and IA Pro, track officer behavior and produce

an electronically generated flag when a threshold set by the administrator is exceeded. As a result, administrators who utilize software tools can rely on the provision of real-time information and no longer count on unsupported suspicion (Walker & Macdonald, 2009).

Limitations involving early intervention systems. As Lersch et al. (2006) highlighted, "Although the use of early intervention systems is becoming more prevalent, the knowledge base of information regarding the actual operation and effectiveness of these systems is minimal" (p. 61). During the review of the literature, numerous studies that delved into the causal roots of police misconduct were noted, yet far fewer were apparent involving the role of early intervention systems in preventing related behaviors. Walker (2001) affirmed this observation by stating that, although numerous questions have been asked regarding the effectiveness of the systems as they gained popularity, little research has been conducted. During one related study, Andre (2007) noted several limitations based upon the fact that, since the early 1980s when the systems were in conceptual status, development has been ongoing through emerging practice. In addition, a lack of literature and corresponding research exists involving specific outcome evaluations as well as the broader examination of officer behaviors and target data monitored through the systems (International Association of Chiefs of Police, 2004). Recognizing the need for further research, participants in a roundtable discussion hosted by the International Association of Chiefs of Police (2004) concluded,

> Effective partnerships between law enforcement leaders and academic researchers are critical to discovering and implementing best police practices. These partnerships are mutually satisfactory: researchers are intensely interested in pursuing such projects, while law enforcement leaders are just as interested in turning the results into enhanced policing practices. (p. 4)

Key elements in EIS evaluations of early intervention systems. Walker (2005) recognized the highly intricate operational aspects of an early intervention system and

explained that LEA administrators must dedicate significant resources involving staffing and funding to ensure proper function of the systems. Even in the early years of use, there are documented examples of system failures as the result of an improper implementation or lackluster attention and maintenance of properly established systems (Walker, 2005).

Obtaining information on the most valuable behaviors to track is an integral part of early identification systems (Walker, 2005). Hughes and Andre (2007) asserted that, although difficult to identify, certain officer behaviors are of great value in predicting misconduct. In a broader scope, Hughes and Andre stated, "For an early identification system to be effective at all, the system must first properly identify the appropriate variables that are causing problems for the agency" (p. 167). However, the general agreement by researchers is that the system should track more pragmatic and universal trends of officers such as resisting-arrest events, use-of-force events, and citizen complaints (Walker et al., 2001).

Regardless of which behaviors LEA administrators believe are most appropriate, the process of continuous evaluation is central to the effectiveness of an early identification system (Lersch et al., 2006). In light of how organizational and environmental influences change over time, Archbold (2005) recommended that administrators consider the need for risk and system evaluation to be perpetual. Bertoia (2008) suggested that a system must be frequently measured within and against the program goals and the mission of the LEA in which it operates in order to acquire needed information for continued improvement. Within such an evaluation, the analysis of the data is critical. As stated by Walker (2005), "It is necessary to continually evaluate the choice of data being collected, how useful they are, whether new data should be collected, and whether the data being collected are of high quality" (p. 42). A

responsibility held by system administrators is to seek and utilize quality, evidence-based data for justifying future decisions and program adjustments, with the intention of their actions to be considered best practice (Regehr & Bober, 2005).

Frank (2009) acknowledged a growing trend for LEA administrators to formally evaluate early identification systems. Noble and Alpert (2009) similarly contended that, without proper methods of analysis, the performance data mean little to law enforcement administrators seeking to determine the effectiveness of a system. The continued cooperation and combined efforts by LEA administrators and the research community could result in such desired, practical results of evaluations (Gibbs & Kendrick, 2011). When addressing the need for outcome evaluations in law enforcement, Bueermann (2012) asserted that the linking of LEA administrators with researchers can produce evidence-based strategies for improved services. During a keynote address, U.S. Associate Attorney General Perrelli (2011) explained that spending public funds on something that does not work is no longer an option.

Studies of early intervention systems. The first study that investigated the use and effectiveness of an early intervention system was conducted by Walker et al. (2001) as part of a National Institute of Justice grant directive. The analysis compared data on officers who received alerts to those who did not. Data and demographics were collected for both groups over a 2-year period before and following system implementation. The purpose of the study was to measure overall agency impact and the quality of system alerts. Results suggested that the system appeared to largely reduce citizen complaints and other targeted behaviors indicative of problematic officer performance. Specifically, Walker et al. (2001) reported a 67% reduction in citizen complaints in Minneapolis, a 62% reduction in citizen complaints in New Orleans, and a change from 4% to 50% of

officers with zero use-of-force complaints in Miami-Dade. However, the report noted limitations such as some LEA administrators claiming to employ a system when, in actuality, the system was either not implemented or not functioning correctly (Walker et al., 2001).

A recent study conducted within the Los Angeles County Sheriff's Department (Bobb, 2009) evaluated the effectiveness of the early identification program and specifically measured the relationships of the individual alert criteria using a multivariate regression model. The researchers reported that the system "performed well, thereby validating the [Los Angeles Sheriff's Department]'s efforts and early identification systems in general" (Bobb, 2009, p. 1).

The Role of Evaluation

Preskill (2008) emphasized the salient role of evaluation by reporting that the public interest in evaluation has increased to epidemic proportion over the past 10 years. Framing evaluation studies as opportunities to learn, Preskill further purported that evaluation professionals should infuse learning in every aspect of theory and practice (p. 129). Froggatt and Hockley (2011) similarly reported that researchers conduct both evaluation studies and applied research to facilitate future decisions and judgments to be made.

Evaluation can play a key role in an organization's strategic planning (U.S. Government Accountability Office, 2012). Although researchers undertake evaluation studies for various reasons that are primarily reflective of stakeholders' needs and requirements of funding sources, related studies generally are designed to answer a range of questions and inform related decisions (Creswell, 2012; Froggatt & Hockley, 2011). Smith et al. additionally reported that evaluation studies are instrumental in (a)

envisioning and predicting the future, (b) preparing effective responses to emerging problems, (c) capitalizing on fresh opportunities, and (d) developing more effective practices. When public resources are used to fund the evaluation object, however, studies are frequently conducted to provide external accountability for the expenditure of funds (U.S. Government Accountability Office, 2012). Along this line of thought, Schwandt (2008) cautioned researchers against affecting public opinion by "being disingenuous and deceptive" (p. 140).

The definitions and applications of evaluation. Researchers subscribe to differing meanings of evaluation. For example, Scriven (1991) defined evaluation as simply a judgment of merit or worth. Incorporating more detail, Mertler and Charles (2010) described evaluation as a judgment of quality and worth of programs, procedures, or materials. Fitzpatrick, Sanders, and Worthen (2010) more succinctly emphasized that evaluation involves a process of identifying and applying specific criteria for determining the evaluation object's merit, value, or worth. Through the use of criteria, the basis of an evaluation becomes clear to stakeholders, and recommendations involving the adoption, continuation, or cessation of the evaluation object are more understandable and justifiable (Fitzpatrick et al., 2010). Fitzpatrick et al. additionally stated that researchers have developed five broad applications of evaluation, based on the evaluation object, within the field of evaluation: (a) product, (b) personnel, (c) program, (d) policy, and (e) performance. With the exclusion of performance valuations, members of the American Evaluation Association (2008) also promoted the evaluation of programs, products, personnel, and policy.

Scriven (as cited in Fitzpatrick et al., 2010) further differentiated the applications of evaluation based on their formative and summative approaches. Fitzpatrick et al.

(2010) emphasized that each approach uniquely affects the choices, decisions, and judgments ensuing from evaluations as well as stakeholders' actions in response to the evaluations. A formative evaluation is designed for two distinct purposes. First, a formative evaluation may be used to identify information for improving the object of the evaluation (Fitzpatrick et al., 2010). As an alternative, a formative evaluation may be conducted to identify the merit, value, or worth of one segment of the object of evaluation (Fitzpatrick et al., 2010). Fitzpatrick et al. differentiated the formative and summative approach to evaluation by stating that a summative evaluation is designed to provide information for making judgments about the adoption, continuation, or cessation of the evaluation object. Scriven (as cited in Gall, Gall, & Borg, 2009) additionally distinguished between the two approaches by stating that formative evaluation is often conducted by the developers, whereas summative evaluation is frequently conducted by individuals beyond the immediate setting to help assure an appropriate element of objectivity. Regardless of approach, the quality of evaluations is critical. In a succinct statement, Chelimsky (1995) asserted, "Telling the truth to people who may not want to hear it is, after all, the chief purpose of evaluation" (p. 54).

The quality of evaluations. The quality of evaluations is dependent on the development of effective research designs (Cooksy & Mark, 2012; U.S. Government Accountability Office, 2012). To this end, representatives of the U.S. Government Accountability Office (2012) identified five specific design steps: (a) clarify the goals and strategies of the evaluation object; (b) develop relevant, effective research questions; (c) select the most appropriate evaluation approach or design for each research question; (d) identify data sources and collection procedures for obtaining credible data to answer each question; and (e) develop analysis approaches to yield valid conclusions relevant to

each question. Cooksy and Mark (2012) underscored the belief that the quality of an evaluation is dependent on three additional factors: (a) the competency of the evaluator, (b) specific aspects of the context or setting where the evaluation is conducted, and (c) "the level and nature of supportive resources that are available in the evaluation community" (p. 79).

Members of the American Evaluation Association (2008) identified and developed five guiding principles in reflection of their mission to "promote ethical practice in the evaluation of programs, products, personnel, and policy" (p. 125). The first principle establishes the expectation that evaluators will conduct systematic, data-based inquiries. This principle is to be accomplished by (a) adhering to high technical standards, (b) examining the shortcomings and strengths of questions and evaluation design with stakeholders, and (c) communicating the components of the evaluation clearly and accurately to stakeholders (American Evaluation Association, 2008, p. 125).

The second principle establishes the expectation that evaluators will provide competent performance to stakeholders (American Evaluation Association, 2008). This principle is to be accomplished by ensuring that the (a) education, skills, and experience of evaluators are commensurate with the parameters of the evaluation; (b) cultural competence of evaluators and the strategies and skills comprising the evaluation are appropriate for use with groups of different cultures; (c) evaluators practice within the limits of their competence; and (d) evaluators maintain and continuously improve their related competencies (American Evaluation Association, 2008).

The third principle establishes the expectation that evaluators will demonstrate honesty and integrity to ensure the honesty and integrity of the evaluation process. This principle is to be accomplished by (a) negotiating honestly with clients regarding all

phases and components of the evaluation, (b) disclosing roles or relationships that may jeopardize the integrity of the evaluation, (c) documenting any changes to the initial evaluation design, (d) explicitly stating how the interests and values of evaluators or stakeholders may relate to the evaluation, (e) accurately reporting procedures and findings of the evaluation, (f) resolving any procedural concerns that may result in misleading findings, and (g) disclosing all sources of financial support involving the evaluation (American Evaluation Association, 2008).

The fourth principle establishes the expectation that evaluators will respect the dignity, security, and self-worth of stakeholders (American Evaluation Association, 2008). This principle is to be accomplished by (a) seeking a comprehensive understanding of the evaluation context; (b) abiding by professional ethics, standards, and regulations involving the protection of participants; (c) maximizing benefits and reducing unnecessary harm that may occur from the evaluation; (d) conducting the evaluation and communicating related findings in a respectful manner; (e) fostering social equity in the collective evaluation processes; and (f) understanding and respecting the culture, religion, gender, disabilities, age, sexual orientation, and ethnicity of stakeholders (American Evaluation Association, 2008).

The fifth, and final, principle establishes the expectation that evaluators will articulate and consider the diversity of general and public interests. This principle is to be accomplished by (a) including relevant perspectives and interests of all stakeholders; (b) collectively considering the immediate outcomes, broad assumptions, implications, and potential side effects of the evaluation; (c) disseminating and providing access of evaluation findings to stakeholders; (d) maintaining a balance between stakeholders' needs and interests; and (e) considering the welfare of society as a whole (American

Evaluation Association, 2008).

Program Evaluation

Urban and Trochim (2009) underscored the importance of program evaluations by stating that program evaluation "is at the heart of efforts to integrate the domains of practice and research" (p. 538). The common practice of not viewing program evaluations from a systems perspective, however, circumvents the potential effectiveness of evaluation and practice (Urban & Trochim, 2009). Viewing program evaluations within a systems perspective, may, at times, encompass the need for evaluators to exercise ethical responsibility and recommend the cessation and closure of programs after completing the evaluations (Eddy & Berry, 2009). Such recommendations are especially challenging when program evaluations are conducted in the public domain, involve public funding, and are simultaneously affected by the political context (Eddy & Berry, 2009).

Fitzpatrick et al. (2010) and Gall et al. (2009) emphasized that a researcher must first identify and clarify the purpose, goals, resources, procedures, and management of a program if the evaluation is to be relevant in determining the program's worth. Gall et al. further emphasized that, whereas many programs have precise purposes and goals, an evaluator must sometimes infer these components. Gall et al. additionally noted that an examination of resources and procedures is important for assisting the researcher in understanding the intended impact of a program. Program procedures, as described by Gall et al., include techniques, strategies, and other collective processes used to accomplish program purposes and goals.

Researchers are in agreement that designing a program evaluation involves three additional steps. The first step is to select the most appropriate evaluation model (Gall

et al., 2009; Mertler & Charles, 2010). The second step is to identify stakeholders (Gall et al., 2009; Mertler & Charles, 2010). Gall et al. (2009) explained that stakeholders may be able to help (a) clarify the need for the evaluation, (b) identify questions for guiding the evaluation, and (c) select the most effective evaluation model. The third step is to identify and analyze the specific components of the program in need of evaluation (Gall et al., 2009; Mertler & Charles, 2010). Brandon and Singh (2009) further emphasized the importance of demonstrating content validity by ensuring that evaluators "explicitly address issues about quality . . . to the extent that study methods are described fully and well" (p. 130).

Program-evaluation models. Gall et al. (2009) underscored the belief that the selected evaluation model must be consistent with the purpose for conducting the evaluation. Models primarily differ in areas involving the (a) purpose of the evaluation; (b) data collection methods; and (c) relationship between the researcher, representatives of the organization, and program administrators (Fitzpatrick et al., 2010). Almost 60 different evaluation models have been identified, forcing the researcher to select the model that is the most effective for each program evaluation (Mertler & Charles, 2010).

Evaluation models are divided into six categories that are based on the approach of each (Fitzpatrick et al., 2010; Mertler & Charles, 2010). As described by Fitzpatrick et al. (2010) and Mertler and Charles (2010), the first is the objectives orientation, which focuses on the program's goals and the extent to which the goals and objectives have been achieved. Mertler and Charles further reported that related evaluations are useful in revising (a) the purposes of the program, (b) specific aspects of the program, and (c) the assessment procedures of the program.

The second category of evaluation models is the management orientation, which

emphasizes the identification and provision of informational needs for those making program-related decisions (Fitzpatrick et al., 2010; Mertler & Charles, 2010). Fitzpatrick et al., as well as Gall et al. (2009), noted that Stufflebeam's CIPP model (Stufflebeam, 2007; Stufflebeam & Shinkfield, 2007) has notably influenced this approach, as the CIPP model was the first to view management, rather than objectives, as the directing force of a program evaluation. Mertler and Charles additionally reported that use of the CIPP model assists program managers by providing a focus for both formative and summative evaluations for use in program development.

The third category of evaluation models is the consumer orientation, which emphasizes the development of accurate, ethical information regarding competing programs or products for the use of consumers (Fitzpatrick et al., 2010; Mertler & Charles, 2010). The strength of this approach is that it reflects and addresses the propriety standard identified by members of the Joint Committee on Standards for Educational Evaluation (2011).

The fourth category of evaluation models is the expertise orientation, which focuses on professional expertise regarding the merit of the program being evaluated (Fitzpatrick et al., 2010; Mertler & Charles, 2010). This approach requires researchers to collect subjective data from experts in the field, justifying concerns that the evaluation findings may be biased (Mertler & Charles, 2010).

The fifth category of evaluation models, which incorporates opposing points of view throughout the evaluation process, is the adversary orientation (Fitzpatrick et al., 2010; Mertler & Charles, 2010). This design purposefully balances the researcher's bias with data from both advocates and adversaries (Mertler & Charles, 2010). Achieving this balance reflects that the model encompasses planned opposition (Mertler & Charles,

2010).

The sixth category of evaluation models is the participant orientation, which incorporates participant involvement to identify values, needs, and data for the evaluation (Fitzpatrick et al., 2010; Mertler & Charles, 2010). Central to this approach is the active involvement of participants through primarily qualitative data collections (Mertler & Charles, 2010). Researchers perceive the participant orientation as credible because of its flexibility, attention to context-based variables, and use of multiple data collection techniques (Mertler & Charles, 2010).

Standards of program evaluation. To achieve and demonstrate relevance, a program evaluation should be structured and systematic (Fitzpatrick et al., 2010; Stufflebeam, 2007; Stufflebeam & Shinkfield, 2007). An effective program evaluation also will produce useful findings that are consistent with many of the 30 evaluation standards identified by members of the Joint Commission on Standards for Educational Evaluation (2011). Under Stufflebeam's leadership, the committee developed criteria, or standards, for program evaluations. The committee's intentions involved improving the professional quality of studies and protecting the evaluation process from those with self-serving motives (Yarbrough, Shulha, Hopson, & Caruthers, 2011).

Committee members initially developed the criteria for an effective program evaluation in 1981 and, over time, revised the criteria through subsequent reviews and publications (Stufflebeam, 2007; Stufflebeam & Shinkfield, 2007; Yarbrough et al., 2011). Although the initial program evaluation standards were limited to four criteria, the current standards include five: (a) utility, (b) feasibility, (c) propriety, (d) accuracy, and (e) evaluation accountability (Joint Committee on Standards for Educational Evaluation, 2011). Yarbrough et al. (2011) emphasized that, collectively, the standards provide

credibility, flexibility, integrity, and validity to both the processes and products of evaluation studies. The utility standards establish expectations pertaining to the usefulness of program evaluations to stakeholders (Joint Committee on Standards for Educational Evaluation, 2011). The standards of feasibility establish expectations concerning the effectiveness and efficiency of program evaluations. The standards of propriety establish expectations pertaining to the practice of ethical standards during the implementation of program evaluations. Accuracy standards establish the expectation that results derived from program evaluations will be dependable and truthful. Standards related to evaluation accountability establish the expectation that adequate documentation will occur during the evaluation process (Joint Committee on Standards for Educational Evaluation, 2011).

The Selected Evaluation Model and Data Collection Techniques

The importance of selecting an effective framework for guiding the development and implementation phases of program evaluations is noted throughout the literature (Adedokun, Childress, & Burgess, 2011; Fitzpatrick et al., 2010; Yarbrough et al., 2011). The CIPP evaluation model, categorized within the management orientation (Stufflebeam, 2007; Stufflebeam & Shinkfield, 2007), is often used to guide program evaluations because of its numerous strengths (Fitzpatrick et al., 2010; Yarbrough et al., 2011). Stufflebeam developed the CIPP model in the 1960s to curtail ongoing corruption in program evaluations and to improve the overall quality and validity of evaluations (Stufflebeam, 2007; Stufflebeam & Shinkfield, 2007). Because the model is focused on eliminating corruption and other flaws, stakeholders can utilize results to improve programs and services (Stufflebeam, 2007; Stufflebeam & Shinkfield, 2007).

As reported by Stufflebeam (2007) and Stufflebeam and Shinkfield (2007), the

CIPP evaluation model (a) is comprehensive, (b) is conducted in an ongoing manner working collaboratively with program administrators and stakeholders, (c) is designed to guide program decisions, and (d) enables the researcher to personally contribute to the evaluation process. Use of the model additionally provides a framework for examining, revising, or improving specific program components (Stufflebeam, 2007; Stufflebeam & Shinkfield, 2007). Coryn, Schröter, and Hanssen (2009) further reported that the CIPP model is effective as a framework for program evaluations conducted within nonprofit settings and public agencies. The model is also widely used in educational settings to examine programs and related guide decision-making processes (Fitzpatrick et al., 2010). Furthermore, the CIPP model is appropriate for examining all aspects of an ongoing program and for guiding both formative and summative evaluations (Stufflebeam, 2007; Stufflebeam & Shinkfield, 2007).

The CIPP evaluation model provides a four-part framework that contributes to the development of specific questions for guiding program evaluations. Context-related evaluations are used to identify the appropriate context, targeted population, problems, and needs assessment as they relate to a specific setting (Stufflebeam, 2007; Stufflebeam & Shinkfield, 2007). Input-related evaluations focus on system capabilities related to resources and strategies for attaining program goals and objectives. Process-related evaluations involve accessing program data for identifying or predicting program defects in the areas of procedural design or program implementation. Finally, product-related evaluations are used to make judgments of outcomes concerning the extent to which program goals and objectives have been achieved (Stufflebeam, 2007; Stufflebeam & Shinkfield, 2007).

Because of the CIPP model's systematic structure, researchers have the

opportunity to objectively explore alternatives for enhancing programs (Stufflebeam, 2007; Stufflebeam & Shinkfield, 2007). Use of the model also assists administrators with four types of decisions: (a) planning, by identifying needs and defining program objectives; (b) structuring, by defining alternative available resources and strategies; (c) implementing, by identifying program quality related to implementation, barriers, and potential revisions for strengthening the program; and (d) recycling, by involving stakeholders in determining the viability of continuing programs (Stufflebeam, 2007; Stufflebeam & Shinkfield, 2007). The CIPP also offers evaluators the ability to be selective in which portions of the model are used based on the client's needs. While explaining such flexibility, Stufflebeam and Shinkfield (2007) stated, "The application of the CIPP model need not always include context, input, process and product evaluations" (p. 362). In light of these collective factors, the CIPP evaluation model is ideal for use in examining programs and supporting leadership decisions regarding systemic programs in both private and public domains. Each of the four segments within the CIPP can be utilized individually to seek information requested by the client (Stufflebeam & Shinkfield, 2007).

After careful consideration of the information sought by the client and the lack of information regarding the development and commencement of the program evaluated (LEA senior executive officer, personal communication, October 18, 2011), the researcher excluded the context and input portions of the Stufflebeam (2003) CIPP model. The researcher utilized the process and product portions of the CIPP as a guide for the evaluation, as they were best suited to provide the information sought by the client of this study (Gairing, 2008).

Rationale for choosing the CIPP evaluation model. An evaluation model that

stresses the importance of eliminating flaws, to include corruption and threats to validity, such as highlighted within the CIPP model, can serve as a valuable tool for stakeholders for use in improving programs and services (Stufflebeam, 2007; Stufflebeam & Shinkfield, 2007). Considered to be of potential value to researchers when seeking information requested by the client is the flexibility offered in the CIPP model, which allows the process and product segments to be utilized individually (Stufflebeam & Shinkfield, 2007). As a result, customizable, targeted information can be examined to support future decisions by leaders at the study site (Stufflebeam & Shinkfield, 2007).

Qualitative techniques used in program evaluation. The use of focus groups during qualitative research has gained popularity in recent years (Sagoe, 2012). Focus groups are often utilized during program evaluations and can be effective for gathering information on programs in all stages of implementation. Focus groups are not designed to result in a concrete consensus but rather for acquiring varying viewpoints and perceptions of those with particular knowledge of the subject (Berger, 2004). Specifically, the use of focus groups during qualitative research enables researchers to collect a large amount of information in a very short amount of time (Wilson & Heinonen, 2011). When dealing with sensitive topics and to quell the concerns of participants in regard to the use of video or audio recordings, researchers may choose to document the focus group interview through handwritten notes (Barnett, 2002).

Survey research also has proven to be a highly valuable tool for institutional researchers (Chen, 2011). Specifically within mixed methods research, the use of surveys can add depth and valuable information to the study (Smith & Stoves, 2012). Survey research has been found to be widely successful in measuring programs and best practices within the law enforcement community (International Association of Chiefs of

Police, 2008). A secure website and research tool created in 1999, SurveyMonkey® (2014), has grown to be a widely utilized tool by researchers within the government arena and is known as a reputable source for the creation and administration of surveys in yielding accurate and valid data.

Quantitative instruments used in program evaluation. *Z* tests are often used to determine the difference between two independent proportions. *Z* tests are often utilized to examine the effectiveness of a program by comparing the before-and-after scores as a tool that "is straightforward and generally yields results easily understood by users" (Jayaratne, Tripodi, & Talsma, 1988, p. 126). Utilizing *z* tests when evaluating proportions often produces results with very high degrees of confidence (Tryon & Lewis, 2009).

In the field of evaluation, methods involving the use of regression and correlation analysis have proven to be powerful and effective (Berger, 2004). "Regression analysis is the statistical methodology of estimating a relationship between a single dependent variable and a set of predictor (explanatory/independent) variables" (Darity, 2008, p. 138). This method has been previously utilized in the field of law enforcement as depicted in a study measuring officer performance and the relationships of various organizational stressors (Shane, 2010). Similarly, when researching correctional staff, this method was utilized to determine the relationships of several predictors of job turnover (Minor, Wells, Angel, & Matz, 2011).

Research Questions

Seven research questions guided this study. Research Questions 1, 2, and 3 addressed the process component of the program evaluation. Research Questions 4, 5, 6, and 7 addressed the product component of the evaluation:

1. What are the perceptions of LEA stakeholders regarding the implementation of the various components of the BTP as it was originally designed?

2. What are the perceptions of LEA stakeholders regarding the sufficiency of staffing and budgetary resources during the implementation of the BTP as it was originally designed?

3. What are the perceptions of LEA stakeholders regarding the sufficiency of education and training received for the implementation and future operation of the BTP as it was originally designed?

4. What are the results relative to incidents of officer misconduct and complaints against officers following implementation of the BTP?

5. What are the results relative to incidents of officer misconduct and complaints against officers following implementation of the BTP when standardizing results by workload?

6. What is the level of statistical significance for each of the BTP alerts in predicting officer misconduct and complaints against officers?

7. What is the level of statistical significance for each of the BTP alerts in predicting officer misconduct and complaints against officers when standardizing results by workload?

Summary

Early intervention systems were developed to limit officer misconduct and complaints against officers, and the BTP was implemented at numerous LEAs for that reason. The literature review provided a foundation for the understanding of officer misconduct and the use of early intervention systems. The literature regarding previous studies of the systems offered ideas in the planning of the BTP evaluation at the LEA

within the research setting. The literature regarding the evaluation of programs and various methods, including strengths and deficiencies of each, assisted the researcher in determining the best suited model. The process and product segments of the CIPP model allowed for specific focus on BTP implementation and effectiveness. When creating a working strategy based on the process and product segments of the CIPP model, the researcher weighed heavily the requests and feedback from LEA stakeholders in an effort to gain insight about effectiveness of the BTP within the research setting; this approach was recommended by Stufflebeam (1999).

In order to answer the research questions of this study, the process and product segments of Stufflebeam's (2003) CIPP model were deemed best suited by the researcher. Specifically, the process segment guided the qualitative evaluation of the BTP implementation, whereas the product segment guided the quantitative evaluation of the BTP effectiveness. The evaluation of the BTP product included overall before- and after-implementation changes, as well as focused analysis of the most effective operational data from within the BTP tracked categories. The combination of answers to the research questions and information from varying perspectives among targeted LEA participants was intended to satisfy the requirements of this dissertation while also providing key information sought by LEA administrators for future decisions regarding the BTP.

Chapter 3: Methodology

Design

This program evaluation, which was conducted over an 8-week period of time, was structured as a mixed methods design. The researcher noted several key points made by authors during the review of related literature that supported the use of this design. Examples in the literature are numerous, yet Creswell (2012) notably explained the complementary nature of mixing qualitative and quantitative methods which can result in a complete set of answers to research questions. Moreover, when considering the focus of this process and product CIPP evaluation, the insight offered by Gall et al. (2009) was weighed heavily and included the defining points of qualitative methodology as being process oriented and quantitative methods as being outcome related. This evaluation incorporated qualitative and quantitative portions concurrently and weighed the respective results equally as they addressed different segments of the evaluation. The researcher triangulated the collected data sources with the intention of relating and connecting the information to expand the knowledge gained through the evaluation. The triangulation of data sources can assist in the development of more complete information and answers to the research questions guiding a study (Creswell, 2012).

To ensure that every reasonable safeguard was taken related to the quality of the study and the trustworthiness of subsequent findings, steps were taken to protect the trustworthiness of data (Chenail, 2011; Mills, 2010). In addition to guarding against the threat of researcher bias, which could affect findings, additional layers of protection included ensuring the confirmability, credibility, dependability, and transferability of data (Chenail, 2011; Mills, 2010). Multiple methods were used to ensure trustworthiness in this study:

1. Credibility in qualitative research is anchored by key indicators that ensure the instruments used in the data collection are appropriate within the context of the research setting and that they are useful in measuring the intended phenomenon (Fitzpatrick et al., 2010). In an effort to enhance and preserve the level of credibility of this study, the researcher piloted both the focus group protocol (see Appendix C) and survey (see Appendix D) with a panel of subject-matter experts. As suggested by Chenail (2011), piloting can result in higher levels of confidence in the instrument tested. In addition, member checks were conducted at the conclusion of the focus group discussions.

2. The transferability of qualitative research results is not often demonstrated in applied research, as most studies are specifically tailored to the topic and research site (Gall et al., 2009; Mills, 2010). However, steps were taken to promote such transferability by providing detailed narratives regarding the qualitative instruments (Merriam, 1998).

3. Confirmability involves the level of objectivity in the analysis and interpretation of results from applied research (Mills, 2010). To enhance the confirmability of results in this study, as suggested by Merriam (1998) and Mills (2010), the researcher utilized several methods and instruments to obtain data from varying perspectives and sources. The researcher then triangulated the collected data by connecting and relating results from several methods in an effort to gain a more complete set of answers to the research questions guiding the study.

4. The dependability of research findings is determined by the validity and stability of collected data (Merriam, 1998; Mills, 2010). As recommended by Mills (2010), the researcher incorporated multiple instruments and sources of data within this study. As recommended by Merriam (1998), the researcher documented all actions

involving data, including collection, analysis, and interpretation, and described such actions in this final dissertation report.

This program evaluation was designed to address the problem that no formal study had been conducted regarding the implementation and effectiveness of the BTP within the LEA within the research setting. The utilization of a mixed methods design resulted in a complete set of answers to the research questions that could only be achieved with the use of both qualitative and quantitative methods (Smith & Stoves, 2012). The evaluation also followed the process and product segments of the CIPP model. The inclusion of both qualitative and quantitative methods in program evaluation has been found to be an effective design by researchers (Greene, Caracelli, & Graham, 1989). Specifically, as stated by Luo and Dappen (2005), a mixed methods evaluation "is not only possible, but more effective, and has higher validity" (p. 109). Stufflebeam (2001) contended that the combined use of qualitative and quantitative methods in program evaluation can answer a wide range of research questions and provide the client with a highly reliable, complete set of results.

As noted by Shaw et al. (2007), "In order to understand 'what works' in crime prevention and criminal justice, evaluations need to focus both on aspects of implementation and outcome" (p. 526). In consideration of the initial requests from the LEA administrators to initiate this study, the CIPP process and product segments were deemed to be the most valuable (LEA senior executive officer, personal communication, October 18, 2011). As expected, the use of these two segments resulted in acquiring the desired answers. This program evaluation provided valuable information that may assist stakeholders in their decision to modify or continue the target program.

Qualitative portion of the design. The researcher used qualitative methods to

evaluate the implementation of the BTP by employing the process element of the CIPP model (Stufflebeam, 2007). As noted in the following Table, the qualitative portion of this study examined the implementation process of the BTP through both a focus group interview of BTP administrators (see Appendix C) and a survey of BTP line supervisors (see Appendix D), designed to acquire input and perceptions from both for comparison and trending from various operational perspectives. As a result of these methods, the researcher collected stakeholder perceptions on the BTP process (Stufflebeam, 2007) in an effort to answer the research questions and thus provide the information sought by LEA administrators.

Table

Alignment of Data Sources With Research Questions

Research Question	Focus Group Discussion	Survey	BTP Data
1	X	X	
2	X	X	
3	X	X	
4			X
5			X
6			X
7			X

Research Question 1 was the following: What are the perceptions of LEA stakeholders regarding the implementation of the various components of the BTP as it was originally designed? To answer this process research question, the researcher included Interview Questions 1 through 4 in the BTP focus group interview (see

Appendix C). In expansion of the effort to gain insight and answer Research Question 1, the researcher included Questions 4 through 7 in the survey of line supervisors (see Appendix D)

Research Question 2 was the following: What are the perceptions of LEA stakeholders regarding the sufficiency of staffing and budgetary resources during the implementation of the BTP as it was originally designed? To answer this process research question, the researcher included Questions 6 through 7 in the focus group interview (see Appendix C). In expansion of the effort to gain insight and answer Research Question 2, the researcher included Questions 1 and 2 in the survey of line supervisors (see Appendix D).

Research Question 3 was the following: What are the perceptions of LEA stakeholders regarding the sufficiency of education and training received for the implementation and future operation of the BTP as it was originally designed? To answer this process research question, the researcher included Question 7 in the focus group interview regarding deficiencies or strengths of the BTP design relative to the sufficiency of training, familiarization with, and introduction of the program (see Appendix C). In expansion of the effort to gain insight and answer Research Question 3, the researcher included Question 3 in the survey of line supervisors regarding the strengths and weaknesses of the BTP as it was originally designed regarding the implementation and functionality for line supervisors (see Appendix D).

Quantitative portion of the design. The researcher used quantitative methods to evaluate the BTP outcomes and effectiveness, as noted in the product segment of the CIPP model (Stufflebeam, 2007), divided into two parts. The portioning consisted of assigning each part two quantitative research questions, with each pair addressed by one

of two very different quantitative methods and approaches. The aforementioned distribution was intended to facilitate a high factor of readability. For the quantitative portion of this study, the researcher collected 6 years of archival data that had been confirmed to be of public record by LEA administrators. The archival data tracked by the BTP, for the population of 918 LEA officers, were relative to officer misconduct and complaints against officers.

In Part 1, to address Research Questions 4 and 5, a single-group, before-and-after approach was taken through the use of a *z* test of proportions. The researcher compared secondary data totals of misconduct and complaints for the population of 918 LEA officers from both before and after implementation of the BTP. This approach resulted in the measurement of overall changes in officer misconduct and complaints against officers subsequent to the implementation of the BTP. Specifically, within Part 1 of the quantitative portion, the researcher compared and contrasted 6 years of data, 2001 through 2003 and 2008 through 2010, representative of periods before and after implementation of the BTP in 2004. These data were used to answer Research Questions 4 and 5 regarding the results relative to incidents of officer misconduct and complaints against officers following implementation of the BTP and, as a second perspective, such incidents when standardizing results by workload.

Further analysis in Part 2 of the quantitative portion included a comprehensive analysis of the most recent 3 years of BTP data, Years 2008 through 2010, pertaining to 238 officers assigned to patrol duties. These data also were used to answer Research Questions 6 and 7 regarding the level of statistical significance for each of the BTP alerts in predicting officer misconduct and complaints against officers and, secondly, in predicting each of the BTP alerts when standardizing results by workload. The researcher

confirmed that the most recent data available for this study were from 2010 because of ongoing open investigations from 2011 and 2012, which were exempt from public record (LEA manager, personal communication, March 31, 2012). The quantitative data collected were of public record.

Within Part 2 of the quantitative portion, the program evaluation followed a nonexperimental, correlational approach through the use of a multivariate regression analysis. The researcher deemed that this approach was best suited to answer Research Questions 6 and 7 to accurately "relate two or more variables to see if they influence each other" (Creswell, 2008, p. 356). This approach resulted in the measurement and identification of which BTP alerts, stemming from each category of BTP tracked data, were most effective in identifying officer misconduct and complaints against officers.

This approach also allowed for the standardization of workload to further enhance confidence in the results. The researcher used this approach with the intent of partially replicating methods used to evaluate an EIS at the Los Angeles Sheriff's Department. In using such methods, the Los Angeles Sheriff's Department researchers answered research questions focused on the degree to which tracked behaviors in the EIS were revealing of complaint and misconduct trends (Bobb, 2009). As noted in a study of the most utilized and effective research methods in the field of law enforcement (Kleck, Tark, & Bellows, 2006), the collection of archival, secondary data and official statistics comprises 58% of all scholarly works and resulted from the use of a nonexperimental design in 96% of the studies.

Following guidelines offered by Creswell (2008), all data held by the LEA regarding the BTP were tracked numerically; hence, a quantitative approach was best suited for the second portion of the study. As also stated by Creswell (2008), it is

important for the researcher to "remember that the problems best suited for quantitative research are those in which trends or explanations need to be made" (p. 62). Consideration was given to the request for an evaluation by the LEA administrator and the audience for whom the researcher was writing (Creswell, 2008). The researcher obtained permission to conduct the study from administrators of the LEA.

Participants

Qualitative. The qualitative portion of the study used two instruments administered to two target populations. The participants for the BTP focus group interview (see Appendix C) included one LEA executive officer, two BTP program managers, and two BTP analysts. These participants were among a very small population at the LEA who held a comprehensive depth of knowledge regarding the implementation and operation of the BTP at the executive level. The participants for the survey of line supervisors (see Appendix D) were 20 supervisors holding the rank of sergeant at the LEA and currently assigned to patrol duties. These participants were among a very small population at the LEA who held a comprehensive depth of knowledge regarding the BTP implementation and functionality at the line level.

The focus group participants were recruited through a confidential, internal memorandum from the researcher to confirm availability. The survey participants were recruited to complete the survey via confidential electronic mail, which included an explanation of the program evaluation, sent through the LEA internal network. The secure website link to the survey, as well as the instructions on how to access and complete the confidential electronic survey, was included in the correspondence.

The researcher included each of the target populations of BTP administrators and line supervisors without the use of sampling. The method of including a target population

without drawing a sample can produce clear, simplified, and expeditious results for presentation to the client (Stufflebeam, 2007). The single exclusion criterion pertained to those employees at the study site who lacked the duty assignment and subsequent potential knowledge to assist in answering the research questions; such employees were outside of the targeted population.

Quantitative. During Part 1 of the quantitative portion of the study, the researcher included all of the approximately archival, deidentified BTP data involving the 918 officers. The purpose of including all of the data was to determine the overall changes in officer misconduct and complaints against officers at the LEA during the years before and after the implementation of the BTP. Although the archived data included officer names and other descriptive information, and were of public record, the researcher did not identify data that could possibly correlate to any participant within the LEA. This method was in consideration of the request by the LEA administrator and was a condition of permission to conduct the evaluation. The data collected consisted of categories of incidents tracked numerically within the BTP without the inclusion of identifying officer data. The data were obtained by the researcher in the form of printed, Microsoft Excel spreadsheets with random numbers assigned to the data set of each officer to facilitate analysis. These random numbers were assigned by the LEA senior executive officer who retained sole knowledge and custody of the method of random numerical assignment.

During Part 2 of the quantitative portion of the study, the researcher used comprehensive, archival, deidentified BTP data pertaining to each of the 238 officers assigned to patrol duties at the LEA during 2008 through 2010. The data had already been obtained during the first quantitative portion, and the researcher extracted the data

for additional analysis. These data were collected in consideration of assertions by LEA administrators that the most extensive data available were for officers assigned to patrol (LEA senior executive officer, personal communication, October 18, 2011). Walker et al. (2001) explained additional reasoning for selecting patrol officers as a study group: "The analysis controlled for assignment to patrol duty on the assumption that citizen complaints against officers and use-of-force incidents are infrequently generated in other assignments" (p. 3).

The researcher obtained and utilized all available, deidentified BTP data involving the target population. Obtaining such comprehensive information without the inclusion of a population sample resulted in a fluid analysis of the data and was conducive to a swift and simplified report for the client of the evaluation (Stufflebeam, 2007). Using these data, the researcher was able to provide LEA administrators with clear results in response to their requested evaluation of the BTP.

Evaluation Model

This evaluation followed the process and product segments of Stufflebeam and Shinkfield's (2007) CIPP model. The CIPP model offers evaluators the ability to be selective in terms of which portions of the model are used based on the needs of clients. While explaining such flexibility, Stufflebeam and Shinkfield stated, "The application of the CIPP model need not always include context, input, process and product evaluations" (p. 362). In light of these collective factors, the CIPP evaluation model is ideal for use in examining programs and supporting leadership decisions regarding systemic programs in both private and public domains.

After careful consideration of the information sought by the client and the lack of information regarding the development and commencement of the program evaluated

(LEA senior executive officer, personal communication, October 18, 2011), the researcher excluded the context and input portions of the Stufflebeam (2003) CIPP model. The researcher used the process and product portions of the CIPP as a guide for the evaluation, as they were best suited for deriving the information sought by the client of this study (Gairing, 2008).

Qualitative portion of the study. During the evaluation of the BTP implementation, the researcher followed the process segment of Stufflebeam's (2003) CIPP evaluation as a guide. Process-related evaluations involve accessing program data for identifying or predicting program defects in the areas of procedural design or program implementation (Stufflebeam, 2007; Stufflebeam & Shinkfield, 2007). The researcher used a focus group interview (see Appendix C) with BTP administrators as well as a survey (see Appendix D) of BTP line supervisors to obtain input, perspectives, and detailed information regarding the implementation of the BTP. Obtaining information on the BTP implementation was of specific interest to LEA administrators (LEA senior executive officer, personal communication, October 18, 2011).

Quantitative portion of the study. During the evaluation of BTP effectiveness, the researcher followed the product segment of Stufflebeam's (2003) CIPP as a guide. Product evaluations are used to make judgments of outcomes concerning the extent to which program goals and objectives have been achieved (Stufflebeam, 2007; Stufflebeam & Shinkfield, 2007). The researcher obtained and analyzed archival data during this portion of the study. The researcher intended to determine both the overall BTP effectiveness as well as which of the BTP tracked data were predictive of officer misconduct and complaints against officers. Obtaining information on the BTP effectiveness was of specific interest to LEA administrators (LEA senior executive

officer, personal communication, October 18, 2011).

Instruments

Qualitative. Two instruments were used in this portion of the study to evaluate the implementation of the BTP at the LEA while following the process segment of Stufflebeam's (2007) CIPP model. The instruments for qualitative data collection were a focus group interview (see Appendix C) and an anonymous, electronic survey (see Appendix D).

When collecting data during the focus group interview, collection was in the form of notes taken by the researcher during the interview. The researcher developed the focus group questions in consideration of stakeholder input during meetings with LEA administrators. In addition, the researcher developed the focus group sessions following criteria and guidelines to ensure the validity and reliability of questions and protocol (Eriksson & Kovalainen, 2008; Sagoe, 2012; Stewart, Shamdasani, & Rook, 2007). When designing focus group questions, a researcher should begin with the most significant of issues being addressed and realize that even a relatively brief set of questions can elicit lengthy discussions (Stewart et al., 2007). In addition, when a researcher is attempting to construct a near-perfect set of focus group questions, it should be taken into consideration that, with permission, modifications can be made during the interview to effectively gain the information sought (Eriksson & Kovalainen, 2008).

Stufflebeam and Shinkfield (2007) suggested that a key element of effective focus group questions is that their design directs the discussion toward the information sought regarding the research questions. Similarly, Fitzpatrick et al. (2010) suggested offering questions that lead the flow of discussion in the direction of the evaluation focus. Following the suggestions by Stufflebeam and Shinkfield, as well as Fitzpatrick et al., the

researcher developed the focus group questions to relate directly to the research questions. In addition, the researcher followed criteria and guidelines designed to ensure the validity and reliability of questions and the protocol (Eriksson & Kovalainen, 2008; Sagoe, 2012; Stewart et al., 2007). Prior to finalizing the focus group interview, the researcher submitted the protocol for review regarding face and content validity by a pilot group of BTP subject-matter experts, which included the LEA chief investigative assistant, an LEA internal affairs assistant, and two information technology executives at the LEA. Based on feedback from the pilot group, minor adjustments were made to the focus group interview questions for better clarity and understanding by participants. Upon completion of the pilot testing, the researcher felt confident that the focus group interview protocol was valid for use in this study.

The electronic survey was drafted by the researcher with the aid of recommendations within SurveyMonkey® (2014), a secure website. When designing the survey (see Appendix D), the researcher drafted original questions based on the research questions to be answered while also considering the relevancy to participants (Iarossi, 2006) with the intent to obtain reliable responses. The use of the survey in an electronic format, accessible via a secure web link, was intended to produce high return rates and quality responses in reflection of convenience and availability for participants (Gay, Mills, & Airasian, 2011). The researcher submitted the survey to a pilot group of subject-matter experts for face and content validation and, upon making the suggested, minor changes to questions for clarity, gained a high level of confidence that the survey was valid for use in this study. The pilot group included the LEA chief investigative assistant, an LEA internal affairs assistant, and two LEA information-technology executives.

Quantitative. The data for the quantitative portion of the study were obtained in

the form of printed Microsoft Excel spreadsheets. Spreadsheets contained deidentified data populated from the BlueTeam software application (Version 3.0) and IA Pro software application (Version 7.0). During Part 1 of the quantitative portion of the study, the researcher determined the overall changes in officer misconduct and complaints at the LEA and collected data of misconduct and complaints involving the entire population of 918 officers at the LEA. These data were from archival, deidentified BTP records provided to the researcher by the LEA administrator for the years before and after the implementation of the BTP at the LEA.

During Part 2 of the quantitative portion of the study, the researcher extracted deidentified BTP data pertaining to each of the 238 officers who were assigned to patrol duties during 2008 through 2010. These data were extracted from the larger population of officer data collected during Part 1 of the quantitative portion of the study. The extracted data for the officers were utilized to examine the individual tracked categories and determine which of the 17 behaviors (see Appendix B) are most valuable as early indicators of officer misconduct and complaints against officers.

Data Collection Procedures

Qualitative data. The data for the qualitative portion of the study were gleaned from both a focus group interview with BTP administrators and a survey of BTP line supervisors (see Appendices C and D). The collection of data during the focus group interview was recorded through handwritten notes documented by the researcher. The focus group interview occurred at the LEA headquarters in the office of the LEA senior executive officer. At the conclusion of the interview, the researcher summarized and repeated responses to group members to ensure the accuracy of handwritten notes. Next, an appointment was made with group members to conduct a member check for accuracy.

All gathered data subsequently were stored in a locked, secure home office of the researcher. The researcher was the only individual with access to the office housing the data.

The data were transcribed into typewritten format within Microsoft Word. The group member check involved offering transcribed and summarized responses for potential clarification or correction. The member check indicated that no changes were necessary, as the documented responses depicted the thoughts and perceptions of focus group members. Finalized, summary responses were documented in Microsoft Word without the use of codes relative to the identification of participants. This procedure, which protected the anonymity of the participants, was recommended by Merriam (1998) who asserted that the establishment of a focus group member's identity is not essential.

When collecting data with the survey of line supervisors (see Appendix D), the researcher used an anonymous electronic survey designed by the researcher with the aid of SurveyMonkey® (2014). The survey remained accessible to participants for 3 weeks. When accessing the electronic survey, participants first saw the participation letter and proceeded to complete the survey by clicking on the acceptance icon. Participants had an unlimited amount of time to complete the survey. If a participant exited the website prior to submission of responses and failed to return to complete the survey, the participant was withdrawn from the study. SurveyMonkey® (2014) did not collect signatures or Internet protocol addresses in order to protect the anonymity of participants. The resulting data collected from the completed electronic surveys were stored for analysis within the secure SurveyMonkey® (2014) website.

Quantitative data. The deidentified data obtained for the quantitative portion of the study were in the form of printed Microsoft Excel computer reports containing data

populated from the LEA BlueTeam (Version 3.0) and IA Pro (Version 7.0) software platforms. The data were provided to the researcher by the LEA senior executive officer. The meeting occurred at the LEA headquarters, within the administration division, in the office of the LEA senior executive officer. All gathered data were stored in a locked, secure home office of the researcher. The researcher was the only individual with access to the office housing the data.

Confidentiality and Custody of Data

The researcher maintained anonymity of all data to "ensure that the identity of subjects cannot be ascertained during the course of the study, in study reports, or in any other way" (Sanders and The Joint Committee on Standards for Educational Evaluation, 1994, p. 203). Ethical standards of Nova Southeastern University, as well at those currently stated by the Joint Committee on Standards for Educational Evaluation (2011), were observed. Once the data were obtained, the researcher stored the data for analysis in a secure home office. The data remained in this secure location throughout the duration of the study and will remain for 3 years after study completion. After 3 years, all documents will be shredded at the study site, placed in garbage bags, and placed in the secure document disposal area of the LEA. All data stored on the computer contained no identifiable information and have been permanently deleted from the computer and all secondary storage sources.

Data Analysis Procedures

Integration of data into Microsoft Word and Excel. Qualitative data were inductively analyzed for themes and categories and documented in narratives using Microsoft Word software. Both the BlueTeam (Version 3.0) and IA Pro (Version 7.0) software tools allowed for the quantitative data to be populated into Microsoft Excel

spreadsheets (Dees, 2003). The information contained in the spreadsheets was then analyzed. Analysis from Excel spreadsheets and Microsoft Word documents is one preferred method in law enforcement research due to the simplicity for describing results to those outside of the research arena and the ability to aggregate data in a format conducive to examining each area of the process (Martinez, 2008).

Qualitative. The qualitative data of the program evaluation were gained from the focus group interview (see Appendix C) and survey (see Appendix D) for documentation in Microsoft Word files. The data from both instruments were inductively analyzed using process, activity, and strategy codes as suggested by Creswell (2008) and Merriam (1998). Further documentation occurred upon the identification of patterns and data groups that were transferred into coded categories. Interpretations were made by the researcher based on gathered data and current literature in an effort to answer the research questions. The findings are reported in this final dissertation report to support the related recommendations. The data obtained during the qualitative portion of the study involved functional details and perceptions of stakeholders, both administrators and line supervisors, regarding the implementation of the BTP.

Quantitative. During the quantitative portion of the program evaluation, two methods of analysis were used. In Part 1 of the quantitative portion, the researcher used a *z* test of proportions; analyses were standardized by workload when comparing total confirmed cases of officer misconduct and complaints against officers from before and after the implementation of the BTP. The data examined in Part 1 of the quantitative portion consisted of LEA totals for staffing, workload, complaints, and officer misconduct. This method measured the total overall effectiveness of the BTP in the reduction of officer misconduct and complaints against officers by comparing data from

before and after BTP implementation. The method was based on the researcher's thorough review of related literature and in consideration of the requested information from LEA administrators (LEA senior executive officer, personal communication, January 7, 2013).

During Part 2 of the quantitative portion of the program evaluation, the researcher utilized a multivariate regression analysis and conducted one regression analysis for each BTP alert type. Based on a thorough review of the related literature and in consideration of the requested information from LEA administrators, the researcher deemed these methods best suited for answering Research Questions 6 and 7. The dependent variables for this analysis were the number of complaints against officers and confirmed acts of officer misconduct. The independent variables for this analysis were the BTP alerts by category. Defining variables was useful in identifying which categories tracked within the BTP were statistically significant in predicting officer misconduct and complaints against officers. This information was originally requested by LEA administrators (LEA senior executive officer, personal communication, January 7, 2013).

Triangulation of qualitative and quantitative data. The researcher related multiple types of data through triangulation in an effort to broaden the understanding gained as a result of the study (Creswell, 2012). When examining the implementation and effectiveness of the BTP, guided by the process and product segments of Stufflebeam's (2003) CIPP, the merging and interrelating of data were imperative for acquiring a comprehensive understanding of each program component. The researcher identified common themes from within the qualitative results of the focus group interview (see Appendix C) and survey (see Appendix D). Furthermore, the researcher was able to interpret results from the quantitative portion which bolstered such themes. The

researcher intended to gain a higher level of validity and confidence in results by including triangulation as part of the design (Creswell & Plano, 2011). In addition to using the mixed methods design combined with triangulation, the researcher acquired the answers to the research questions and other vital information requested by LEA administrators.

Summary

This program evaluation was guided by the process and product segments of Stufflebeam's (2003) CIPP model of evaluation and utilized a mixed methods design (see Figure). By conducting the evaluation, the researcher was able to answer the three qualitative research questions regarding the process segment and the four research questions involving the product segment.

During the qualitative portion of the study, the perceptions of stakeholders regarding the various components of the BTP, sufficiency of staffing, budget, and training for effective implementation were acquired. The quantitative portion of the study evaluated BTP effectiveness in reducing incidents of officer misconduct and complaints against officers and also provided an indication of which alerts were valid indicators of such trends.

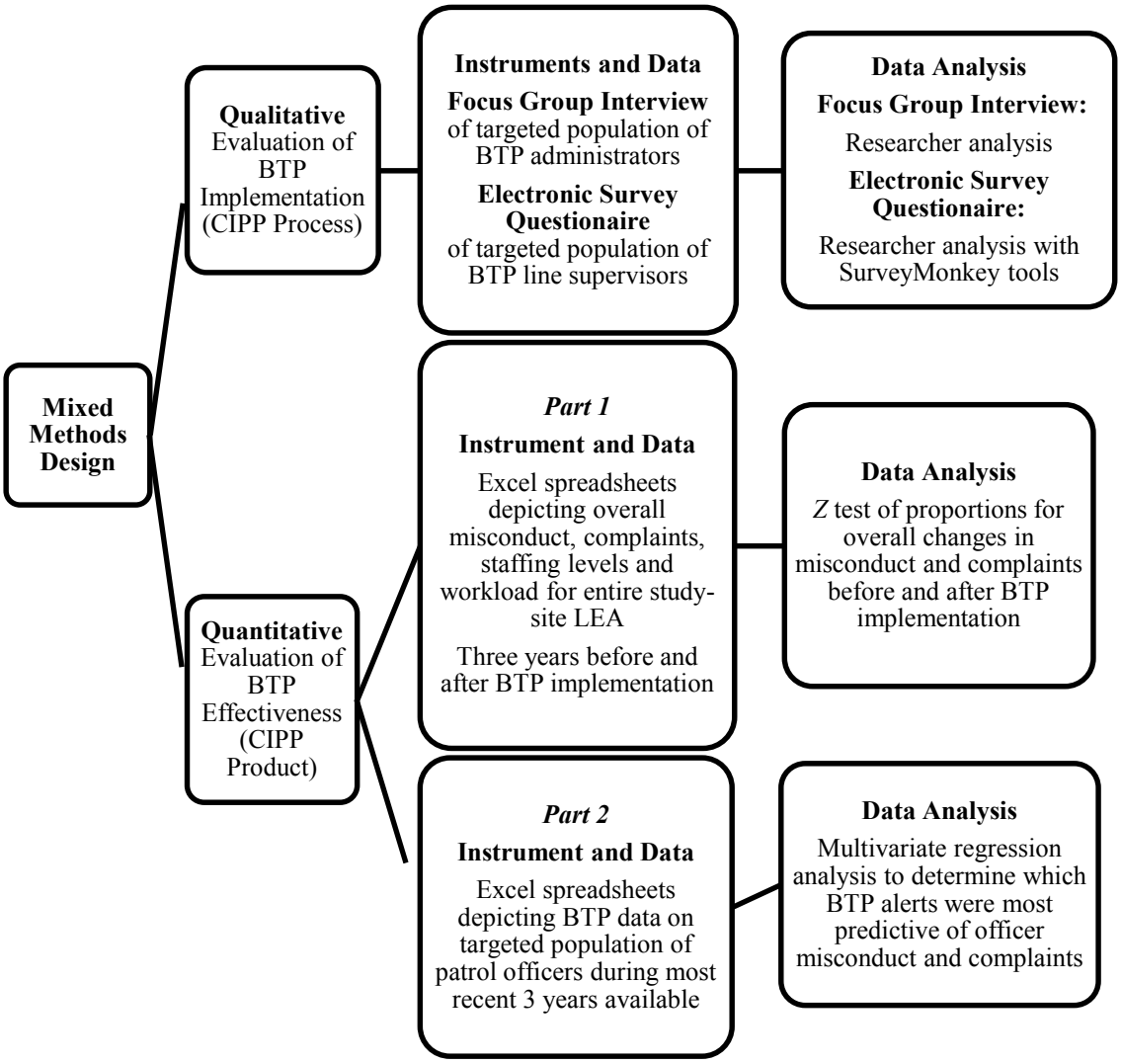

Figure. Research design informational chart. BTP = LEA BlueTeam (Version 3.0) and IA Pro (Version 7.0); CIPP = context, input, process, and product (Stufflebeam, 2003; Stufflebeam & Shinkfield, 2007); LEA = law enforcement agency.

Chapter 4: Results

The problem addressed through this program evaluation, which was guided by the process and product components of the CIPP model (Stufflebeam, 2007; Stufflebeam & Shinkfield, 2007) and incorporated a mixed methods design, was that no formal study had been conducted regarding the implementation and effectiveness of the BTP within the LEA within the research setting. The purpose of the evaluation was to determine the effectiveness of the BTP implementation and operation regarding the reduction of officer misconduct and complaints against officers at the LEA within the research setting. The evaluation was guided by seven research questions. Research Questions 1, 2, and 3 addressed the process component of the program evaluation; Questions 4, 5, 6, and 7 addressed the product component of the evaluation:

1. What are the perceptions of LEA stakeholders regarding the implementation of the various components of the BTP as it was originally designed?

2. What are the perceptions of LEA stakeholders regarding the sufficiency of staffing and budgetary resources during the implementation of the BTP as it was originally designed?

3. What are the perceptions of LEA stakeholders regarding the sufficiency of education and training received for the implementation and future operation of the BTP as it was originally designed?

4. What are the results relative to incidents of officer misconduct and complaints against officers following implementation of the BTP?

5. What are the results relative to incidents of officer misconduct and complaints against officers following implementation of the BTP when standardizing results by workload?

6. What is the level of statistical significance for each of the BTP alerts in predicting officer misconduct and complaints against officers?

7. What is the level of statistical significance for each of the BTP alerts in predicting officer misconduct and complaints against officers when standardizing results by workload?

This chapter presents the results derived from the data collection and subsequent analyses. Results are presented first for the process component and then for the product component. Within each section, results are organized by research question.

Process Component

The implementation of the BTP. Research Question 1 was as follows: What are the perceptions of LEA stakeholders regarding the implementation of the various components of the BTP as it was originally designed? This question was answered through the analysis of responses to Questions 1 through 4 of the focus group interview and Questions 4 through 7 of the survey. Results are presented separately for each question of the data collection and then the research question is answered as the conclusion of this section.

Question 1 of the focus group interview (see Appendix C) was: Was the BTP implemented as designed regarding anticipated compatibility and functionality with current agency technology? In response to this question, the agreement of all five participants reflected the theme that the implementation was a success regarding overall functionality and compatibility. This perception was supported by statements indicating that the needed technology was already in place prior to implementation. Participants additionally identified initial concerns prior to the implementation of the BTP involving the compatibility and functionality of the BTP. One focus group participant described the

concerns by stating the following:

> Over the years the LEA has suffered many technology setbacks and challenges due to a lack of funding for equipment and skilled technology staff. So as we are understandably a little gun-shy, we moved into the implementation phase with a very conservative, reasonable mindset of expected success regarding the compatibility and functionality of the BTP within our technology platforms and operations.

Question 2 of the focus group interview (see Appendix C) was: Was the BTP implemented as designed regarding tracked data categories intended to detect trends of officer misconduct? Two themes were identified that reflected the collective responses of all participants. First, agreement was expressed involving the successful implementation of tracked data categories regarding the detection of officer-misconduct trends. Second, all participants expressed the previous hope of including officer-assignment criteria and officer-workload criteria in tracked data categories. Prior to implementing the BTP, these changes would have required major technology upgrades and were deemed too costly for their allotted budgetary resources. In summary, one participant stated, "I think we were successful in implementing the tracked categories regarding officer misconduct which we ultimately selected based on the resources available at the time."

Question 3 of the focus group interview (see Appendix C) was: Was the BTP implemented as designed regarding tracked data categories intended to detect trends of complaints against officers? Two themes were identified that reflected the collective responses of all participants. First, agreement was expressed involving the successful implementation of tracked data categories, as planned, regarding the detection of complaints against officers. Secondly, participants expressed agreement that the previous hope of including officer- assignment criteria and officer-workload criteria in tracked data categories would have required major technology upgrades and was ultimately

deemed too costly within their allotted budgetary resources. Similar to statements made in response to the related focus group Question 2, one participant stated, “I feel we succeeded in implementing the tracked officer data regarding complaints partially as a result of maximizing our resources drawn from a very challenging budget and manpower allocation.”

Question 4 of the focus group interview (see Appendix C) was: What were the anticipated and unanticipated events or challenges that took place during the implementation of the BTP? In response to this question, the agreement of all five participants reflected the theme involving the reiteration of the previous hope to include officer-workload and officer-assignment criteria within the tracked data categories. All participants additionally agreed on the challenges involving the unanticipated expense of required technology upgrades and the subsequent unanticipated result of not including these categories. As one participant stated, “During difficult budget years it was not cost efficient to purchase the upgrades, yet it was certainly not in the best interest of the agency to delay the implementation of the BTP until such funds were available.”

Question 4 of the survey (see Appendix D) was: What is your perception of the BTP regarding the functionality and compatibility with current agency technology such as hardware and software platforms? Two themes were identified from participant responses. First, the majority expressed positive views regarding end-user functionality and compatibility with current agency technology. Second, all participants expressed satisfaction with the frequency of system outages. Fifteen of the 20 participants (75%) believed the BTP was relatively easy to use and made statements such as “not bad,” “user friendly,” and “clear and straightforward.” The remaining five participants provided statements describing the BTP visual experience such as “nighttime view is lacking” and

"difficult to use when working in the dark during the night shift."

Question 5 of the survey (see Appendix D) was: What is your perception of the BTP functionality regarding the tracked data categories intended to detect trends of officer misconduct? Two themes were identified that reflected the responses of participants. First, participants were affirmative regarding BTP functionality within the tracked data categories intended to detect trends of officer misconduct. Second, participants expressed agreement involving the functionality of real-time updates pertaining to the detection of officer misconduct. Twelve of the 20 participants (60%) included positive mention of the live updates of data categories offering supervisors immediate access to accurate numbers. One participant further stated, "The program and categories seem to be very functional and provide updated data almost immediately upon entry which can serve to be valuable when reviewing an officer's most recent data."

Question 6 of the survey (see Appendix D) was: What is your perception of the BTP functionality regarding the tracked data categories intended to detect trends of complaints against officers? Two themes were identified in the analysis of responses. First, participants were affirmative regarding BTP functionality within the tracked data categories intended to detect trends of complaints against officers. Second, participants expressed agreement involving the functionality of real-time updates pertaining to the detection of complaints against officers. In concurrence with responses to the related Survey Question 5, all participants offered positive responses regarding the functionality of the BTP tracked data intended to detect trends of complaints against officers. Also similar to the aforementioned responses, 11 of the 20 participants (55%) mentioned the value of live updates and data accuracy. This trend is depicted by one participant who included, "It is a very useful tool when you have just received a complaint against an

officer and you can count on the data to be accurate and current."

Question 7 of the survey (see Appendix D) was: What is your perception of the BTP pertaining to ease of use or challenges in overall functionality for line supervisors? Two themes were identified in responses; the first was the overall ease of access and functionality of the BTP for line supervisors. The second theme was the noted exception of nighttime visibility. All participants offered positive responses regarding the overall ease of access and functionality for line supervisors. One example offered by a respondent was, "In our line of work it is seldom that we encounter a program such as the BTP which is accurately designed for use in the field – the BTP is overall very accessible and functional." Sixteen of the 20 participants (80%), however, included the mention of the BTP being challenging to use while working at night because of dim lighting. One participant depicted this concern by stating, "When we utilize computers in the patrol vehicles at night, programs need to have a night view mode to reduce lighter and brighter colors which decrease our night vision."

Research Question 1 was as follows: What are the perceptions of LEA stakeholders regarding the implementation of the various components of the BTP as it was originally designed? The analysis of participant responses indicated six distinct themes:

1. The implementation of the BTP was a success regarding overall functionality, in part, because of the real-time updates. Stakeholders collectively described the ability of the program to track data categories involving officer-misconduct trends as well as officer-assignment criteria and officer-workload criteria within tracked data.

2. The implementation of the BTP was a success regarding compatibility with agency technology.

3. The BTP was relatively easy to use.

4. The frequency of system outages was satisfactory.

5. Implementation of the BTP fulfilled objectives that would have otherwise required major technology upgrades that exceeded budgetary resources.

6. The one drawback of the BTP was the limited nighttime visibility.

Staffing and Budgetary Resources for the BTP. Research Question 2 was as follows: What are the perceptions of LEA stakeholders regarding the sufficiency of staffing and budgetary resources during the implementation of the BTP as it was originally designed? This question was answered through the analysis of responses to Questions 5 through 7 of the focus group interview and Questions 1 and 2 of the survey. Results are presented separately for each question of the data collection.

Question 5 of the focus group interview (see Appendix C) was: What are your thoughts about the sufficiency of budget for appropriate BTP implementation, particularly in regard to staffing and ancillary resources? One theme was identified in the responses; participants agreed that the budget was sufficient in terms of staffing and ancillary resources. Participants further agreed that the resources were appropriate during the years that were challenging in the areas of fiscal responsibility and manpower allocation. One participant expounded on this theme by stating, “I think we all worked with extreme focus, in an efficient manner, and kept things as simple as possible. All of which would have only been more complicated if we were allowed cart blanch with our planning.”

Question 6 of the focus group interview (see Appendix C) was: What are your thoughts about the sufficiency of workload allocation and time for appropriate BTP implementation? Two themes were identified in responses. First, the implementation of

the BTP was timely; second, the workload was appropriately allocated. In support of these themes, agreement was expressed that administrators provided both a sufficient workload allocation as part of their normal duties as well as a reasonable amount of dedicated time to the BTP implementation. All participants also agreed with one participant who stated, "Both workload and time allocations prompted us to work in a focused and efficient manner yet were not so limited to cause rushed work and the potential of resulting errors."

Question 7 of the focus group interview (see Appendix C) was: What deficiencies or strengths do you perceive regarding the design of the BTP relative to the sufficiency of training, familiarization with, and introduction of the program? Two themes were identified in the analysis of responses. First, participants perceived that they had received thorough training in the use of the BTP. Second, participants reported receiving sufficient support as they began using the program. All five participants concurred that the BTP software provider, as well as the technology consultants, provided thorough training as well as support during the planning and implementation phases. Participants referenced the provision of personal and telephone-based support as well as the opportunity to see support through electronic mail. Participants agreed with the statement made by one participant as follows: "We could not have been so successful without the initial training and ongoing patience during continued step-by-step assistance provided by both the software provider as well as technology consultants."

Question 1 of the survey (see Appendix D) was: What is your understanding of the purpose for the implementation and use of the BTP as originally designed by the LEA administrators? Two themes were identified in participant responses. First, participants expressed agreement that the purpose of the BTP was to reduce officer misconduct.

Second, participants were in agreement that the BTP was implemented to reduce complaints against officers. All 20 participants indicated that administrators were clear in explaining the purpose of the BTP prior to implementation.

Question 2 of the survey (see Appendix D) was: What is your opinion regarding the sufficiency of training and education previously provided by LEA administrators on the topic of BTP purpose and use for line supervisors? Responses were useful in identifying two themes. The first was that the initial BTP training was sufficient. The second theme was that the appropriate resources were provided for answering questions. Sixteen of the 20 participants (80%) indicated that LEA administrators provided both comprehensive training and appropriate channels for questions via telephone or electronic mail. One participant mentioned, "Even though the initial training was thorough, many of us had follow-up questions which were promptly addressed by the BTP staff." Alternatively, four participants (20%) stated that, although the initial training was sufficient, they worked the night shift and were unable to receive immediate assistance by telephone during their overnight work hours. All four participants conceded, however, that they were able to receive support via electronic mail on the next business day.

Research Question 2 was as follows: What are the perceptions of LEA stakeholders regarding the sufficiency of staffing and budgetary resources during the implementation of the BTP as it was originally designed? The analysis of participant responses indicated six distinct themes:

1. The LEA budgetary resources, including staffing and ancillary resources, were sufficient for implementing the BTP.

2. The implementation of the BTP was timely.

3. The workload directly attributable to the BTP was appropriately allocated.

4. The necessary training for the implementation of the BTP was provided. Furthermore, stakeholders perceived the usability of the program was increased because of the quality of the training.

5. The needed support was provided as stakeholders began using the BTP; consequently, the usability of the program was increased. Support was provided through personal, telephone, and electronic-mail responses. The only related concern was that support was not available during the overnight work hours.

6. Stakeholders were aware that the BTP was implemented to reduce officer misconduct and complaints against officers.

Education and Training for the BTP. Research Question 3 was as follows: What are the perceptions of LEA stakeholders regarding the sufficiency of education and training received for the implementation and future operation of the BTP as it was originally designed? This question was answered through the analysis of responses to Question 7 of the focus group interview and Question 3 of the survey. Results are presented separately for each question of the data collection.

Question 7 of the focus group interview (see Appendix C) was: What deficiencies or strengths do you perceive regarding the design of the BTP relative to the sufficiency of training, familiarization with, and introduction of the program? Responses to this question were also used to answer Research Question 2. In review, two themes were identified in the analysis of responses. First, participants perceived that they had received thorough training in the use of the BTP. Second, participants reported receiving sufficient support as they began using the program. All five participants concurred that the BTP software provider, as well as the technology consultants, provided thorough training as well as support during the planning and implementation phases.

Question 3 of the survey (see Appendix D) was: What is your opinion regarding the strengths and weaknesses of the BTP as it was originally designed regarding the implementation and functionality for line supervisors? Two themes were identified in responses of the participants. First, all agreed that the ease of access to both functional and accurate information was the most significant strength of the BTP. As one participant offered, "It is a good tool to utilize as a supervisor when seeking accurate and up-to-date information about an officer in a simple format." The second theme was identified by 16 of the 20 participants (80%). These individuals noted the lack of sufficient viewing in low lighting as the most significant weakness of the BTP. The remaining four participants provided no response depicting weaknesses of the program.

Research Question 3 was as follows: What are the perceptions of LEA stakeholders regarding the sufficiency of education and training received for the implementation and future operation of the BTP as it was originally designed? Participant responses supported two themes that were initially identified in the analysis of responses to Research Question 2. First, the training was sufficient for using the BTP; second, ample support was provided as stakeholders began using the BTP. Responses additionally supported two themes initially identified in the analysis of responses to Research Question 1. First, the primary strength of the BTP was the ease of access to functional information. Second, the limited nighttime visibility of the BTP was a salient drawback to the program. Participant responses additionally indicated the theme that the provision of accurate information was a primary strength of the BTP.

Product Component

Incidents of officer misconduct and complaints against officers. Research Question 4 was as follows: What are the results relative to incidents of officer misconduct

and complaints against officers following implementation of the BTP? This question was answered through the analysis of archival, deidentified BTP data pertaining to 918 officers. The total number of misconduct cases, complaints, and members were calculated for the pre- and post-time period. Then, utilizing the *z* test, the proportions of each were tested for significant differences between the two periods. The following results were noted in the analyses (see Tables E1 and E2 in Appendix E):

1. The proportion of misconduct to total complaints for the 3 years leading up to the intervention was 0.130, while the same proportion 3 years after the intervention fell to 0.127. The difference was not statistically significant at $p < 0.05$ level of significance.

2. The proportion of misconduct to total members for the 3 years leading up to the intervention was 0.030, while the same proportion 3 years after the intervention fell to 0.021. The difference was statistically significant at $p < 0.01$ level of significance.

3. The proportion of total complaints to total members for the 3 years leading up to the intervention was 0.229, while the same proportion 3 years after the intervention fell to 0.163. The difference was statistically significant at $p < 0.01$ level of significance.

Research Question 4 was as follows: What are the results relative to incidents of officer misconduct and complaints against officers following implementation of the BTP? Analysis of the data indicated that the proportion of (a) misconduct to total complaints was not statistically significant at $p < 0.05$ level of significance, (b) misconduct to total members was statistically significant at $p < 0.01$ level of significance, and (c) total complaints to total members was statistically significant at $p < 0.01$ level of significance (see Tables E1 and E2 in Appendix E).

Incidents of officer misconduct and complaints against officers by workload. Research Question 5 was as follows: What are the results relative to incidents of officer

misconduct and complaints against officers following implementation of the BTP when standardizing results by workload? This question was answered through the analysis of archival, deidentified BTP data pertaining to 918 officers utilizing the *z* test of proportions. The proportions of misconduct to total workload and complaints against officers to total workload were tested, utilizing the *z* test, for significant differences between the pre-post time periods.

Two results were noted in the analyses (see Tables E1 and E2 in Appendix E). First, the proportion of misconduct to total workload for the 3 years leading up to the intervention was 0.000077, while the same proportion 3 years after the intervention fell to 0.000051. The difference was statistically significant at $p < 0.01$ level of significance. Second, the proportion of total complaints to total workload for the 3 years leading up to the intervention was 0.00059, while the same proportion 3 years after the intervention fell to 0.000401. The difference was statistically significant at $p < 0.001$ level of significance.

Research Question 5 was as follows: What are the results relative to incidents of officer misconduct and complaints against officers following implementation of the BTP when standardizing results by workload? Analysis of the data indicated that the proportion of misconduct to total workload was statistically significant at $p < 0.01$ level of significance and the proportion of total complaints to total workload was statistically significant at $p < 0.001$ level of significance (see Tables E1 and E2 in Appendix E).

Predicting officer misconduct and complaints against officers. Research Question 6 was as follows: What is the level of statistical significance for each of the BTP alerts in predicting officer misconduct and complaints against officers? To answer this question, the researcher used archival, deidentified BTP data on each of the 238 officers assigned to patrol duties at the LEA during 2008 through 2010. The data

necessary for answering this research question were obtained during the first quantitative portion of the analysis and used in answering Research Questions 4 and 5. The applicable data for the target population were extracted for use in conducting descriptive statistics and multivariate regression analyses.

As reflected in Table E3 in Appendix E, descriptive statistics for Years 2008 through 2010 reflect that mean calls, or workload, for each participant were 3,716 (*SD* = 2,374). Complaint categories included citizen complaints (mean = 0.82, *SD* = 1.23), agency complaints (mean = 0.10, *SD* = 0.37), employee complaints (mean = 0.02, *SD* = 0.17), inmate complaints (mean = 0.11, *SD* = 0.37), and use-of-force complaints (mean = 0.23, *SD* = 0.50). Predictor variables included chain of command (mean = 0.04, *SD* = 0.23), response to resistance (mean = 3.27, *SD* = 4.80), Taser (mean = 0.66, *SC* = 1.38), resist arrest without violence (mean = 4.46, *SD* = 6.48), resist arrest with violence (mean = 0.69, *SD* = 1.15), failure to appear (mean = 0.12, *SD* = 0.34), vehicle pursuit (mean = 0.57, *SD* = 1.01), and vehicle accidents (mean = 0.90, *SD* = 1.03). Dependent variables were sustained misconduct (mean = 0.16, *SD* = 0.50) and total complaints (mean = 1.27, *SD* = 1.68).

A multiple regression analysis was conducted to determine whether statistically significant relationships existed between the variables of citizen complaints, agency complaints, employee complaints, inmate complaints, use-of-force complaints, chain of command, response to resistance, Taser, resist arrest without violence, resist arrest with violence, failure to appear, vehicle pursuits, vehicle accidents, and the dependent variable of sustained misconduct for Years 2008 through 2010 (see Table E4 in Appendix E).

The R^2 value for the model identified within Table E4 (see Appendix E) was low ($R^2 = 0.32$). About 32% of the variation in sustained misconduct was explained by this set

of predictor variables. The possibility of multicollinearity was also checked using the variance inflation factor (VIF) diagnostic. Results reflected that VIF values ranged from 1.05 to 4.33; therefore, no problems with multicollinearity were detected. The assumption of normality of the residuals was checked using a histogram of the residuals and upheld based on a visual assessment of the histogram.

Two of the independent variables were statistically related to sustained misconduct. These two variables were agency complaints ($B = 0.72$, $t = 9.59$, $p < 0.01$), and Taser ($B = -0.06$, $t = -2.06$, $p < 0.05$). As agency complaints increased, the number of sustained misconducts also increased significantly. Also, as Taser increased, sustained misconducts decreased significantly (see Table E4 in Appendix E).

A multiple regression analysis was then conducted to determine whether statistically significant relationships existed between the variables of chain of command, response to resistance, Taser, resist arrest without violence, resist arrest with violence, failure to appear, vehicle pursuits, and vehicle accidents and the dependent variable of total complaints for Years 2008 through 2010 (see Table E5 in Appendix E).

The R^2 value for the model identified within Table E5 (see Appendix E) was low ($R^2 = 0.29$). About 29% of the variation in total complaints was explained by this set of predictor variables. The possibility of multicollinearity was also checked using the VIF diagnostic. Results reflected that VIF values ranged from 1.05 to 4.16; therefore, no problems with multicollinearity were detected. The assumption of normality of the residuals was checked using a histogram of the residuals and upheld based on a visual assessment of the histogram.

Two of the independent variables were statistically related to total complaints. These two variables were chain of command ($B = 1.01$, $t = 2.43$, $p < 0.05$) and resist

arrest without violence ($B = 0.10$, $t = 3.94$, $p < 0.01$). As chain of command increased, the number of total complaints also increased significantly. Also, as resist arrest without violence increased, the total number of complaints increased significantly (see Table E5 in Appendix E).

Research Question 6 was as follows: What is the level of statistical significance for each of the BTP alerts in predicting officer misconduct and complaints against officers? Analysis of the data resulted in the following four findings:

1. The independent variable of agency complaints was statistically related to sustained misconduct. As agency complaints increased, the number of sustained misconducts also increased significantly (see Table E4 in Appendix E).

2. The independent variable of Taser was statistically related to sustained misconduct. As Taser increased, sustained misconducts decreased significantly (see Table E4 in Appendix E).

3. The independent variable of chain of command was statistically related to total complaints. As chain of command increased, the number of total complaints also increased significantly (see Table E5 in Appendix E).

4. The independent variable of resist arrest without violence was statistically related to total complaints. As resist arrest without violence increased, the total number of complaints increased significantly (see Table E5 in Appendix E).

Predicting officer misconduct and complaints against officers when standardizing by workload. Research Question 7 was as follows: What is the level of statistical significance for each of the BTP alerts in predicting officer misconduct and complaints against officers when standardizing results by workload? To answer this question, the researcher used archival, deidentified BTP data on each of the 238 officers

assigned to patrol duties at the LEA during 2008 through 2010. The data were obtained during the first quantitative portion of the analysis and used in answering Research Questions 4, 5, and 6. The applicable data necessary for answering this research question were extracted for use in conducting descriptive statistics and multivariate regression analyses (see Table E3 in Appendix E).

A multiple regression analysis was conducted to determine whether statistically significant relationships existed between citizen complaints, agency complaints, employee complaints, inmate complaints, use-of-force complaints, chain of command, response to resistance, Taser, resist arrest without violence, resist arrest with violence, failure to appear, vehicle pursuits, vehicle accidents, and the dependent variable of sustained misconduct. The variable of workload, based on total calls, for Years 2008 through 2010 was included in the regression analysis.

The R^2 value for the model identified within Table E6 (see Appendix E) was low ($R^2 = 0.32$). About 32% of the variation in sustained misconduct was explained by this set of predictor variables. The possibility of multicollinearity was also checked using the VIF diagnostic. Results reflected that VIF values ranged from 1.04 to 4.30; therefore, no problems with multicollinearity were detected. The assumption of normality of the residuals was checked using a histogram of the residuals and upheld based on a visual assessment of the histogram.

Two of the independent variables were statistically related to sustained misconduct. These two variables were agency complaints ($B = 0.73$, $t = 9.60$, $p < 0.01$) and Taser ($B = 0.06$, $t = -2.06$, $p < 0.05$). As agency complaints increased, the number of sustained misconducts also increased significantly. Additionally, as Taser increased, sustained misconducts decreased significantly (see Table E6 in Appendix E).

A multiple regression analysis was then conducted to determine whether statistically significant relationships existed between the independent variables of chain of command, response to resistance, Taser, resist arrest without violence, resist arrest with violence, failure to appear, vehicle pursuits, vehicle accidents, and the dependent variable of total complaints. The variable of workload, based on total calls, for Years 2008 through 2010 was included in the regression analysis.

The R^2 value for the model identified within Table E7 (see Appendix E) was low ($R^2 = 0.29$). About 29% of the variation in total complaints was explained by this set of predictor variables. The possibility of multicollinearity was also checked using the VIF diagnostic. Results reflected that VIF values ranged from 1.05 to 4.13; therefore, no problems with multicollinearity were detected. The assumption of normality of the residuals was checked using a histogram of the residuals and upheld based on a visual assessment of the histogram.

Two of the independent variables were statistically related to total complaints. These two variables were chain of command ($B = 1.01$, $t = 2.42$, $p < 0.05$) and resist arrest without violence ($B = 0.10$, $t = 3.98$, $p < 0.01$). Also, as resist arrest without violence increased, the number of total complaints also increased significantly (see Table E7 in Appendix E).

Research Question 7 was as follows: What is the level of statistical significance for each of the BTP alerts in predicting officer misconduct and complaints against officers when standardizing results by workload? Analysis of the data resulted in the following four findings:

1. As also noted when not standardizing results by workload, the independent variable of agency complaints was statistically related to sustained misconduct. As

agency complaints increased, the number of sustained misconducts also increased significantly (see Table E6 in Appendix E).

2. As also noted when not standardizing results by workload, the independent variable of Taser was statistically related to sustained misconduct. As Taser increased, sustained misconducts decreased significantly (see Table E6 in Appendix E).

3. As also noted when not standardizing results by workload, the independent variable of chain of command was statistically related to total complaints. As chain of command increased, the number of total complaints also increased significantly (see Table E7 in Appendix E).

4. As also noted when not standardizing results by workload, the independent variable of resist arrest without violence was statistically related to total complaints. As resist arrest without violence increased, the total number of complaints increased significantly (see Table E7 in Appendix E).

Chapter 5: Discussion

This concluding chapter begins with an overview of the program evaluation conducted on behalf of administrators of the LEA within the research setting. An elaboration and interpretation of results are then provided. Conclusions, drawn from the analysis of results, are presented next. After identifying the limitations of the study, recommendations for program improvement are presented. The final section concludes with additional recommendations for future research.

Overview of the Study

This program evaluation was designed to determine the effectiveness of the BTP implementation and operation regarding the reduction of officer misconduct and complaints against officers at the LEA within the research setting. The program was implemented during 2004 and was in its 10th year of implementation at the time the evaluation was conducted. Central to the investigation were comprehensive analyses of the most recent 3 years of data, 2008 through 2010. This investigation was the initial evaluation of the program. As reflected in the following text, the research was guided by seven research questions within the process and product components of the CIPP evaluation model (Stufflebeam, 2007; Stufflebeam & Shinkfield, 2007).

Process components. Qualitative data were acquired from (a) one LEA executive officer, (b) two BTP program managers, (c) two BTP analysts, and (d) 20 supervisors holding the rank of sergeant at the LEA and currently assigned to patrol duties to answer the following three questions:

1. What are the perceptions of LEA stakeholders regarding the implementation of the various components of the BTP as it was originally designed?

2. What are the perceptions of LEA stakeholders regarding the sufficiency of

staffing and budgetary resources during the implementation of the BTP as it was originally designed?

3. What are the perceptions of LEA stakeholders regarding the sufficiency of education and training received for the implementation and future operation of the BTP as it was originally designed?

Product component. Quantitative BTP data for the two periods of 2001 through 2003 and 2008 through 2010, pertaining to 918 officers, were used to answer Research Questions 4 and 5. Question 4 was as follows: What are the results relative to incidents of officer misconduct and complaints against officers following implementation of the BTP? Question 5 was: What are the results relative to incidents of officer misconduct and complaints against officers following implementation of the BTP when standardizing results by workload?

A subset of the data used to answer Research Questions 4 and 5 that pertained to 238 officers assigned to patrol duties at the LEA during 2008 through 2010 at the LEA was used to answer Research Questions 6 and 7. Question 6 was: What is the level of statistical significance for each of the BTP alerts in predicting officer misconduct and complaints against officers? Question 7 was: What is the level of statistical significance for each of the BTP alerts in predicting officer misconduct and complaints against officers when standardizing results by workload?

Elaboration and Interpretation of Results

Process component of the evaluation. The first three research questions were useful in exploring stakeholder perceptions regarding the implementation of the BTP pertaining to the (a) various components of the BTP, (b) sufficiency of staffing and budgetary resources of the BTP, and (c) sufficiency of education and training for the

implementation and future operation of the BTP. Through the collection and analysis of pertinent data, three overarching themes were identified. First, the BTP was perceived as successful regarding overall functionality, compatibility with agency technology, ease of use, accuracy, and fiscal resources. Second, the workload related to the BTP was appropriately allocated, and the quality of initial training and subsequent support enabled stakeholders to gain intended benefits from the program in a timely manner. Third, the limited nighttime visibility of the BTP was a salient concern warranting continued consideration.

Product component of the evaluation. The remaining four research questions were useful in guiding the statistical analysis of performance data pertaining to the incidents of officer misconduct and complaints against officers following implementation of the BTP and the level of statistical significance for each of the BTP alerts in predicting officer misconduct and complaints against officers. In both areas, the researcher examined and compared the data both with and without standardizing results by workload, with workload being represented by total calls. Through the collection and analysis of the data, six relationships were identified:

1. When comparing incidents of officer misconduct and complaints against officers, the proportions of misconduct to total members and total complaints to total members were statistically significant when not standardizing by workload (see Tables E1 and E2 in Appendix E).

2. When comparing incidents of officer misconduct and complaints against officers, the proportions of misconduct to total workload and total complaints to total workload were statistically significant when standardizing by workload (see Tables E1 and E2 in Appendix E).

3. When comparing the predictability of BTP alerts with incidents of officer misconduct and complaints against officers, the relationship between the independent variable of agency complaints and the dependent variable of sustained misconduct was statistically significant both when and when not standardizing results by workload (see Tables E4 and E6 in Appendix E).

4. When comparing the predictability of BTP alerts with incidents of officer misconduct and complaints against officers, the relationship between the independent variable of Taser usage and the dependent variable of sustained misconduct was statistically significant both when and when not standardizing results by workload (see Tables E4 and E6 in Appendix E).

5. When comparing the predictability of BTP alerts with incidents of officer misconduct and complaints against officers, the relationship between the independent variable of chain of command and the dependent variable of total complaints was statistically significant both when and when not standardizing results by workload (see Tables E5 and E7 in Appendix E).

6. When comparing the predictability of BTP alerts with incidents of officer misconduct and complaints against officers, the relationship between the independent variable of resist arrest without violence and the dependent variable of total complaints was statistically significant both when and when not standardizing results by workload (see Tables E5 and E7 in Appendix E). The inclusion of resist arrest without violence is key for an effective BTP, as it is a viable indicator of subsequent complaints (Walker et al., 2001).

Conclusions

Prior to this program evaluation, no formal study had been conducted within the

research setting regarding the implementation and effectiveness of the BTP. By conducting the evaluation, the researcher was able to examine both perceptual and statistical data to determine the effectiveness of the BTP implementation and operation regarding the reduction of officer misconduct and complaints against officers at the LEA. A review of the professional literature endorsed the importance of evaluating early intervention systems in an effort to ensure that instances of officer misconduct and complaints against officers are reduced (Braga & Schnell, 2013; Police Executive Research Forum, 2013; Weisburd & Neyroud, 2011; Worden et al., 2013). Braga (2010) additionally cited the need for researchers to partner with police agencies to help ensure the credibility of findings from related investigations. Also according to Braga, the methodology used in this program evaluation was the most effective approach for ensuring internal validity.

The belief also exists, however, that a lack of literature involving the evaluation of early intervention systems exists and that each evaluation is unique to the respective research setting (Bazley, Mieczkowski, & Lersch, 2009; Worden et al., 2013). One primary reason for the limited generalizability of research findings is that tracked performance data vary among agencies (Harris, 2012). Kane and White (2009) additionally underscored the belief that numerous definitions of police misconduct exist and that it is difficult to apply specific definitions to actual cases of misconduct. Additional discussion of the literature in comparison with conclusions derived from the evaluation regarding the process and product components is presented in the following text.

Process component. The collective perceptions of participants reflect that stakeholders of the LEA within the research setting view the BTP as an effective early

intervention system for identifying incidents of officer misconduct and complaints against officers. The belief that early intervention systems, regardless of program developer and manufacturer, are effective in achieving the intended purpose of reducing incidences of misconduct is consistent with the literature (Harris, 2009, 2012; Hassell & Archbold, 2010; Kane & White, 2009). Along with other innovative programs such as police immersion within communities, the use of early intervention systems is supported as a proactive approach for coping with the common budget reductions made necessary by the national recession involving efforts to maintain high standards of accountability (Harris, 2009; Spinelli, 2010; Weisburd & Neyroud, 2011).

Furthermore, stakeholder support for utilizing the BTP is reflective of the ongoing changes in policies and practices currently reforming policing across the United States (Weisburd & Neyroud, 2011). Participants additionally highlighted the quality of system training and ongoing support as the reason for gaining intended benefits from the program. The literature cited these factors as essential elements for the successful implementation of early intervention systems (Kane & White, 2009; Police Executive Research Forum, 2013; Weisburd & Neyroud, 2011).

The nighttime visibility of the BTP is the single limitation of the system identified by participants of this program evaluation. Because law enforcement is a 24-hour service, this limitation must be addressed. Bazley et al. (2009) emphasized that the substance of early intervention systems varies within the profession and, furthermore, that no universal model or set of performance indicators has been developed. The focus of the literature is on which behaviors to monitor and how to apply system results to either reprimand officers or provide behavioral interventions through expanded training (Harris, 2012). Nighttime visibility, however, is a shortcoming of the BTP that cannot be overlooked.

Product component. The data analyses central to this program evaluation indicate that the BTP is effective as an effective early intervention system for both identifying and predicting incidents of officer misconduct and complaints against officers. The findings are consistent with the literature (Bazley et al., 2009; Braga & Schnell, 2013; Harris, 2009, 2012; Hassell & Archbold, 2010). Assuming the system limitation involving nighttime vision can be effectively addressed, the remaining question would be how to best use the system for identifying police officers in need of supportive interventions or reprimands.

The question of how to best utilize an early intervention system was underscored by Bazley et al. (2009) and Harris (2012). The extended review of the literature was helpful not in answering this question but in identifying the many parameters with which the question might be answered. After reviewing the current professional literature, this researcher concluded that the question might be answered based upon the two approaches of officer characteristics and the effectiveness with which the early intervention system is used.

In general, patterns of officer misconduct align with those of criminal behavior; related behaviors are initially exhibited and then may sustain, decrease, or cease over time as the individual matures (DeLisi & Piquero, 2011; Harris, 2009, 2012). The possibility of ongoing changes in conduct contributes to the difficulty of identifying the officers who may be most likely to engage in misconduct or receive agency or citizen complaints (Harris, 2009, 2012). Harmon (2009) additionally stated that police misconduct is neither accidental nor inevitable but occurs as a result of systemic deficiencies, supporting the belief that the fault may be shared among officers and administrators.

Several researchers, such as Hassell and Archbold (2010) and Harris (2009, 2012), have adopted what DeLisi and Piquero (2011) described as a "life-course perspective" to misconduct and complaints involving misconduct (p. 289). The review of the literature suggested numerous possible characteristics that may place police officers at greater risk of receiving citizen complaints; some of these characteristics also place officers at greater risk of engaging in misconduct.

The first characteristic involves gender. Female officers, when compared to their male counterparts, are less likely to be subjected to citizen complaints involving misconduct (Hassell & Archbold, 2010). Furthermore, male officers are more likely to receive numerous complaints within a short time period (Hassell & Archbold, 2010). Some studies indicate that cultural and ethnic minorities, when compared to Caucasians, are more likely to receive a higher number of complaints of misconduct (Hassell & Archbold, 2010).

Age is a third characteristic that may place police officers at greater risk of receiving citizen complaints; young police officers, in comparison with older officers, are more likely to receive more complaints involving misconduct (Harris, 2009; 2012; Hassell & Archbold, 2010). Experience as police officers is another characteristic that may place police officers at greater risk of receiving citizen complaints; officers with less police experience than their peers are more likely to receive a greater percentage of complaints involving misconduct (Hassell & Archbold, 2010). This characteristic was also identified by Harris (2012) who further stated that, while police offers are in their first few years of service, they are more likely to engage in misconduct than their experienced peers. The propensity to engage in misconduct, however, is also noted among officers during the latter portion of their police careers (C. J. Harris, 2011; C.

Harris, 2011; Stinson, Liederbach, & Freiburger, 2010).

The level of education is another characteristic that may place police officers at greater risk of receiving citizen complaints. After conducting a review of the research, Hassell and Archbold (2010) reported that officers without a college education are more likely to receive more complaints involving misconduct. Psychological factors may also place police officers at greater risk of receiving citizen complaints (DeLisi & Piquero, 2011; Harris, 2012). Hassell and Archbold (2010) further reported that officers who are more aggressive, or who issue a larger percentage of citations and arrest more suspects, often receive a greater percentage of complaints involving misconduct. Included within this group are the officers who interact with and question suspects upon arrest (Hassell & Archbold, 2010).

Also included within the group of more aggressive officers are those who are younger and less experienced who typically "do more to detect crime" by initiating citizen contacts, actively patrol, and record crime reports at a higher rate than more experienced peers (Harris, 2009, p. 194). These final points cast doubt on the use of total workload, or calls, as a determinant factor of misconduct or complaints against officers although this approach is recommended in the literature (Harris, 2012). An additional concern is that the indicators and thresholds necessary for implementing early intervention systems may deter the more aggressive officers from being as productive as they may otherwise be (Worden et al., 2013).

The second approach that may assist in determining how to best utilize an early intervention system involves the effectiveness with which the system is used. Bazley et al. (2009) emphasized that, although tracking the frequency involving the use of force is the most common indicator, merely tracking the number of times force is used is not

effective in identifying officers in need of further support or discipline. Numerous beliefs were noted in the review of the literature, yet the salient factor is that no specific approach has been universally adopted. The three primary decisions that must be made involve (a) ascertaining the way in which the early intervention system will be used, (b) identifying the indicators of misconduct that will be tracked, and (c) determining the threshold at which the system will issue an alert.

Ascertaining the way in which the early intervention system will be used is the first primary decision that must be made. The literature rarely included a discussion of how to apply the information acquired through early intervention systems. Harris (2009) suggested that, while most systems are used to track misconduct, their use typically involves behaviors that would not warrant termination. Recognizing this practice, representatives of the Police Executive Research Forum (2013) further suggested changing the focus of the systems to one of overall performance improvement by tracking collective data for both misconduct and consistent performance of duties without misconduct. In this way, the systems could be used for commendations as well as intervention or disciplinary action.

Identifying the indicators of misconduct that will be tracked is the second primary decision that must be made. Representatives of the Police Executive Research Forum (2013) recommending that, at a minimum, the following three indicators be used: (a) the improper use of force, (b) unlawful stops and searches, and (c) biased policing. One of these three recommended indictors was used in this present program evaluation.

Several other indicators of misconduct, including the improper use of force as identified by representatives of the Police Executive Research Forum (2013), were identified and used in this present program evaluation, yet others exist. For example,

although this evaluation did not include off-duty misconduct, this is one area that can be considered as an indicator (Bazley et al., 2009; Hassell & Archbold, 2010). Other indicators, such as corruption and consumption or trafficking of drugs, also could be included (Harris, 2009). Bazley et al. (2009) additionally suggested that LEA administrators broaden the scope of misconduct indicators to include those that are merely "suggestive of misconduct and/or a need for some type of departmental help or assistance" (p. 109). Furthermore, current trends indicate that the number of misconduct indicators is increasing over time and includes a wider range of misconduct that may involve the failure to not perform assigned duties (Bazley et al., 2009). Again, the dearth in the literature places LEA administrators and practitioners at a disadvantage when determining which indicators are more likely to assist in pinpointing the officers in the most need of support or disciplinary actions.

Determining the threshold at which the system will issue an alert is the third primary decision that must be made. Because of the dearth in the literature, Harris (2012) suggested removing all thresholds and placing responsibility of analyzing officer behavior patterns on LEA supervisors. As an alternative, Harris additionally suggested combining categories of indicators to determine an overall score rather than to identify a threshold using only one indicator. To more definitively determine the threshold for issuing an alert, Harris, as well as Worden et al. (2013), recommended the ongoing evaluation of the program by analyzing patterns of behaviors involving shift and assignment data as well as peer norms. Budgetary limitations, however, may preclude the use of ongoing evaluations. An additional recommendation made by Worden et al. is to consider workload, or total calls, when determining thresholds in order to avoid false positives that can ultimately impact staff morale.

The current program evaluation was effective in determining the effectiveness of the BTP implementation and operation regarding the reduction of officer misconduct and complaints against officers at the LEA within the research setting. Results supported the continued use of the system providing the limited nighttime vision can be resolved. Analyzing results derived from the investigation and comparing findings with the literature, however, underscores the need for continued study not only within the LEA serving as the research setting but throughout the United States and within the international domain as well. As previously noted, this investigation was an initial attempt within the research setting to examine the BTP, but continued study to improve the effective application of the system within the LEA is warranted.

Limitations

Three limitations may have affected either the validity or trustworthiness of findings derived from this program evaluation, yet the researcher attempted to minimize each:

1. The internal validity of the investigation may have been affected regarding the minimal amount of secondary data and number of similar early intervention programs for comparison. This possibility was identified by Stufflebeam and Shinkfield (2007) as inherent to program evaluations involving understudied topics.

2. As anticipated, results of this program evaluation may be specific to the LEA within the research setting and may not be generalizable to other agencies because of the differences in identified indicators and thresholds of officer misconduct and complaints against officers. This limitation was previously noted by Walker et al. (2001) when explaining the difficulty in comparing early intervention systems when each agency may have employed various other tools to reduce officer misconduct and enhance

performance.

3. Potential limitations during the focus group interview included the reluctance of participants to provide honest and forthcoming responses because of the confidential nature of the law enforcement community.

Recommendations for Program Improvement

Based on findings of this program evaluation, the BTP is an effective early intervention system for use in the LEA serving as the research setting. The one area of needed improvement, which is of critical concern, is the low quality of nighttime vision. It is recommended that this shortcoming be addressed to determine possible solutions for improving this program component.

Recommendations for Future Research

Based on findings of this program evaluation, and supported by the extensive review of the literature, the researcher recommends future research involving the application of early intervention systems. There is a continued deficiency of research in the literature regarding early intervention systems as well as evaluation studies in the field of law enforcement (Bazley et al., 2009; Worden et al., 2013). Initial determinations, as well as formative evaluations, must occur involving the following three areas: (a) ascertaining the way in which the early intervention system will be used, (b) identifying the indicators of misconduct that will be tracked, and (c) determining the threshold at which the system will issue an alert. As Harmon (2009) accentuated, identifying and minimizing the systemic contributors of officer misconduct and complaints against officers requires "structurally changing police departments to create accountability for officers and supervisors and foster norms of professional integrity" (p. 1). By creating accountability for officers and supervisors and fostering professional

integrity, the agency will continually progress toward needed reform within the profession.

References

Adedokun, O. A., Childress, A. L., & Burgess, W. D. (2011). Testing conceptual frameworks of nonexperimental program evaluation designs using structural equation modeling. *American Journal of Evaluation, 32*, 480-493. http://dx.doi.org/10.1177/1098214011401368

Alpert, G. P., & Noble, J. J. (2009). Lies, true lies, and conscious deception: Police officers and the truth. *Police Quarterly, 12*, 237-254. http://dx.doi.org/10.1177/1098611108327315

American Evaluation Association. (2008). Guiding principles for evaluators. *American Journal of Evaluation, 29*, 125-126. http://dx.doi.org/10.1177/1098214008029002061

Andre, L. B. (2007). *Police misconduct and early warning systems: An exploratory examination of influential variables* (Master's thesis). Available from ProQuest Dissertations and Theses database. (UMI No. 1452126)

Archbold, C. A. (2005). Managing the bottom line: Risk management in policing. *Policing: An International Journal of Police Strategies and Management, 28*, 30-48. http://dx.doi.org/10.1108/13639510510580968

Associated Press. (2005, October 29). New Orleans officers fired for leaving posts. *The Boston Globe.* Retrieved from http://www.boston.com/news/nation/articles/2005/10/29/new_orleans_officers_fired_for_leaving_posts/

Barnett, J. M. (2002). *Focus groups tips for beginners* (TCALL occasional research report No. 1). Retrieved from the Texas Center for Adult Literacy & Learning website: http://www-tcall.tamu.edu/orp/orp1.htm

Barrick, M. R., & Mount, M. K. (2012). Nature and use of personality in selection. In N. Schmitt (Ed.), *The Oxford handbook of personnel assessment and selection* (pp. 225-251). New York, NY: Oxford University Press.

Bazley, T. D., Mieczkowski, T., & Lersch, K. M. (2009). Early intervention program criteria: Evaluating officer use of force. *Justice Quarterly, 26*, 107-124. http://dx.doi.org/10.1080/07418820801989742

Berger, D. E. (2004). Using regression analysis. In J. S. Wholey, H. P. Hatry, & K. E. Newcomer (Eds.), *Handbook of practical program evaluation* (2nd ed., pp. 479-505). San Francisco, CA: Jossey-Bass.

Bertoia, T. (2008). *Developing an early intervention system for police misconduct in a law enforcement agency* (Research and Issue Paper No. 1). Sydney, Australia: NSW Police Integrity Commission.

Bobb, M. J. (2009). *The Los Angeles County Sheriff's Department 27th semiannual report.* Los Angeles, CA: Police Assessment Resource Center.

Braga, A. A. (2010). Setting a higher standard for the evaluation of problem-oriented policing initiatives. *Criminology & Public Policy*, *9*, 173–182. http://dx.doi.org/10.1111/j.1745-9133.2010.00618.x

Braga, A. A., & Schnell, C. (2013). Evaluating place-based policing strategies: Lessons learned from the Smart Policing Initiative in Boston. *Police Quarterly, 16*, 339-347. http://dx.doi.org/10.1177/1098611113497046

Brandon, P. R., & Singh, J. M. (2009). The strength of the methodological warrants for the findings of research on program evaluation use. *American Journal of Evaluation, 30*, 123-157. http://dx.doi.org/10.1177/1098214009334507

Bueermann, J. (2012). Being smart on crime with evidence-based policing. *NIJ Journal, 269*. Retrieved from http://www.nij.gov/journals/269/evidence.htm

Caillouet, B. A., Boccaccini, M. T., Varela, J. G., Davis, R. D., & Rostow, C. D. (2010). Predictive validity of the MMPI-2 PSY-5 Scales and facets for law enforcement officer employment outcomes. *Criminal Justice and Behavior, 37*, 217-238. http://dx.doi.org/10.1177/0093854809351948

Chelimsky, E. (1995). Comments on the guiding principles. *New Directions for Program Evaluation, 1995*(66), 53-54.

Chen, P.-S. D. (2011). Finding quality responses: The problem of low-quality survey responses and its impact on accountability measures. *Research in Higher Education, 52*, 659-674. http://dx.doi.org/10.1007/s11162-011-9217-4

Chenail, R. J. (2011). Interviewing the investigator: Strategies for addressing instrumentation and researcher bias concerns in qualitative research. *The Qualitative Report, 16*, 255-262. Retrieved from http://www.nova.edu/ssss/QR/QR16-1/interviewing.pdf

CI Technologies. (2010a). BlueTeam (Version 3) [Computer software]. Bellingham, WA: Author.

CI Technologies. (2010b). IA Pro (Version 7) [Computer software]. Bellingham, WA: Author.

Commission on Accreditation for Law Enforcement Agencies. (2010). *Law enforcement program: Eligibility criteria*. Retrieved from http://www.calea.org/content/law-enforcement-program-eligibility-criteria

Cooksy, L. J., & Mark, M. M. (2012). Influences on evaluation quality. *American Journal of Evaluation, 33*, 79-87. http://dx.doi.org/10.1177/1098214011426470

Coryn, C. L. S., Schröter, D. C., & Hanssen, C. E. (2009). Adding a time-series design element to the success case method to improve methodological rigor: An application for nonprofit program evaluation. *American Journal of Evaluation, 30*, 89-92. http://dx.doi.org/10.1177/1098214008326557

Creswell, J. W. (2008). *Educational research: Planning, conducting, and evaluating quantitative and qualitative research* (3rd ed.). Upper Saddle River, NJ: Pearson/Merrill Prentice Hall.

Creswell, J. W. (2012). *Educational research: Planning, conducting, and evaluating quantitative and qualitative research* (4th ed.). Boston, MA: Pearson.

Creswell, J. W., & Plano, C. V. L. (2011). *Designing and conducting mixed methods research* (2nd ed.). Los Angeles, CA: Sage.

Cuttler, M. J., & Muchinsky, P. M. (2006). Prediction of law enforcement training performance and dysfunctional job performance with general mental ability, personality, and life history variables. *Criminal Justice and Behavior, 33,* 3-25. http://dx.doi.org/10.1177/0093854805282291

Darity, W. A. (Ed.). (2008). *International encyclopedia of the social sciences* (2nd ed.). Independence, KY: Cengage Learning.

Dees, T. (2003). Up close: Internal affairs: Internal affairs, management software. *Law and Order, 51*(5), 88-94.

Dekmar, L. M. (2010). Handling citizen complaints through proactive methodology. *The Police Chief, 77*(4), 50-54.

DeLisi, M., & Piquero, A. R. (2011). New frontiers in criminal careers research, 2000-2011: A state-of-the-art review. *Journal of Criminal Justice, 30*, 289-301. http://dx.doi.org/10.1016/j.jcrimjus.2011.05.001

Eddy, R. M., & Berry, T. (2009). The evaluator's role in recommending program closure: A model for decision making and professional responsibility. *American Journal of Evaluation, 30*, 363-376. http://dx.doi.org/10.1177/1098214009339931

Eriksson, P., & Kovalainen, A. (2008). *Qualitative methods in business research*. Thousand Oaks, CA: Sage.

Federal Bureau of Investigation. (2010). *Operation guard shack*. Retrieved from http://www.fbi.gov/news/stories/2010/october/operation-guard-shack/operation-guard-shack

Fitch, B. D. (2011, October). Focus on ethics: Rethinking ethics in law enforcement. *FBI Law Enforcement Bulletin*. Retrieved http://www.fbi.gov/stats-services/publications/law-enforcement-bulletin/october-2011/focus-on-ethics

Fitzpatrick, J. L., Sanders, J. R., & Worthen, B. R. (2010). *Program evaluation: Alternative approaches and practical guidelines* (4th ed.). Upper Saddle River, NJ: Pearson.

Florida Department of Law Enforcement. (2009). *FDLE bulletin*. Tallahassee, FL: Author.

Florida State University. (2000). *Florida Law Enforcement Research Coalition—Final report*. Tallahassee, FL: Author.

Forero, C. G., Gallardo-Pujol, D., Maydeu-Olivares, A., & Andrés-Pueyo, A. (2009). A longitudinal model for predicting performance of police officers using personality and behavioral data. *Criminal Justice and Behavior, 36*, 591-606. http://dx.doi .org/10.1177/0093854809333406

Frank, J. (Ed.). (2009). Conceptual, methodological, and policy considerations in the study of police misconduct. *Criminology and Public Policy, 8,* 733-736. http://dx .doi.org/10.1111/j.1745-9133.2009.00590.x

Froggatt, K., & Hockley, J. (2011). Action research in palliative care: Defining an evaluation methodology. *Palliative Medicine, 25*, 782-787. http://dx.doi .org/10.1177/0269216311420483

Gaffigan, S. J., & McDonald, P. P. (1997). *Police integrity: Public service with honor* (NCJ 163811). Washington, DC: U.S. Department of Justice.

Gairing, A. L. (2008). *Long-term outcome study of at-risk youth following private out-of-home therapeutic intervention* (Unpublished doctoral dissertation). Nova Southeastern University, Fort Lauderdale, FL.

Gall, M., Gall, J., & Borg, W. (2009). *Educational research: An introduction* (8th ed.). New York, NY: Longman.

Gay, L. R., Mills, G. E., & Airasian, P. W. (2011). *Educational research: Competencies for analysis and applications* (10th ed.). New York, NY: Pearson.

Gibbs, M., & Kendrick, C. (2011). *San Diego Police Department: Enhancing cultures of integrity. Building law enforcement early intervention systems technical assistance guide.* Washington, DC: U.S. Department of Justice Office of Community Oriented Policing Services.

Goldstein, H. (1977). *Policing a free society*. Cambridge, MA: Ballinger.

Gottschalk, P. (2011). Management challenges in law enforcement: The case of police misconduct and crime. *International Journal of Law and Management, 53*, 169-181. http://dx.doi.org/10.1108/17542431111133409

Greene, J. C., Caracelli, V. J., & Graham, W. F. (1989). Toward a conceptual framework for mixed-method evaluation designs. *Educational Evaluation and Policy Analysis, 11,* 255-274. Retrieved from http://www.jstor.org/stable/1163620

Grogan, B. (2011, October). Perspective: Leadership moments. *FBI Law Enforcement Bulletin*. Retrieved from http://www.fbi.gov/stats-services/publications/law-enforcement-bulletin/october-2011/perspective

Harmon, R. (2009). Promoting civil rights through proactive policing reform. *Stanford Law Review, 62*, 1-68. Retrieved from http://www.stanfordlawreview.org/

Harris, C. (2011). Problem behaviors in the later portion of officers' careers. *Policing: An International Journal of Police Strategies and Management, 34*, 135-152. http://dx.doi.org/10.1108/13639511111106650#sthash.UBadzr9b.dpuf

Harris, C. J. (2009). Exploring the relationship between experience and problem behaviors: A longitudinal analysis of officers from a large cohort. *Police Quarterly, 12*, 192-213. http://dx.doi.org/10.1177/1098611108327314

Harris, C. J. (2010a). *Pathways of police misconduct: Problem behavior patterns and trajectories from two cohorts*. Durham, NC: Carolina Academic Press.

Harris, C. J. (2010b). Problem officers? Analyzing problem behavior patterns from a large cohort. *Journal of Criminal Justice, 38*, 216-225. http://dx.doi.org/10.1016/j.jcrimjus.2010.01.003

Harris, C. J. (2011). The relationship between career pathways of internal and citizen complaints. *Police Quarterly, 14*, 142-165. http://dx.doi.org/10.1177/1098611111404141

Harris, C. J. (2012). The residual career patterns of police misconduct. *Journal of Criminal Justice, 40,* 323-332. http://dx.doi.org/10.1016/j.jcrimjus.2012.06.010

Hassell, K. D., & Archbold, C. A. (2010). Widening the scope on complaints of police misconduct. *Policing: An International Journal of Police Strategies and Management, 33*, 473-489. http://dx.doi.org/10.1108/13639511011066863

Hickman, M. J., & Piquero, A. R. (2009). Organizational, administrative, and environmental correlates of complaints about police use of force: Does minority representation matter? *Crime and Delinquency, 55*, 3-27. http://dx.doi.org/10.1177/0011128708316977

Huberts, L. W. J. C., Kaptein, M., & Lasthuizen, K. (2007). A study of the impact of three leadership styles on integrity violations committed by police officers. *Policing: An International Journal of Police Strategies and Management, 30*, 587-607. http://dx.doi.org/10.1108/13639510710833884

Hughes, F., & Andre, L. (2007). Problem officer variables and early-warning systems. *The Police Chief, 74*(10), 164-172.

Iarossi, G. (2006). *The power of survey design: A user's guide for managing surveys, interpreting results, and influencing respondents*. Washington, DC: World Bank.

International Association of Chiefs of Police. (2004). *Unresolved problems and powerful potentials: Improving partnerships between law enforcement leaders and university based researchers.* Alexandria, VA: Author.

International Association of Chiefs of Police. (2008). *Improving 21st century policing through priority research: The IACP's national law enforcement research agenda*. Alexandria, VA: Author.

Jayaratne, S., Tripodi, T., & Talsma, E. (1988). The comparative analysis and aggregation of single-case data. *The Journal of Applied Behavioral Science, 24,* 119-128. http://dx.doi.org/10.1177/0021886388241012

Jiao, A. Y. (2010). Controlling corruption and misconduct: A comparative examination of police practices in Hong Kong and New York. *Asian Journal of Criminology, 5*, 27-44. http://dx.doi.org/10.1007/s11417-009-9077-4

Johnson, B. R., & Bridgmon, P. B. (2009). Depriving civil rights: An exploration of 18 U.S.C. 242 criminal prosecutions 2001–2006. Criminal *Justice Review, 34*, 173-195. http://dx.doi.org/10.1177/0734016808326220

Joint Committee on Standards for Educational Evaluation. (2011). *Program evaluation standards statements*. Retrieved from http://www.jcsee.org/program-evaluation-standards-statements

Kane, R. J., & White, M. D. (2009). Bad cops. *Criminology and Public Policy, 8*, 737-769. http://dx.doi.org/10.1111/j.1745-9133.2009.00591.x

Kaplan, P. J. (2009). Looking through the gaps: A critical approach to the LAPD's rampart scandal. *Social Justice, 36*(1), 61-81.

Kinnaird, B. A. (2008). *Exploring police liability profiles: Proximate cause analysis of police misconduct in the San Francisco, CA Police Department*. Saarbrücken, Germany: VDM Verlag.

Kleck, G., Tark, J., & Bellows, J. J. (2006). What methods are most frequently used in research in criminology and criminal justice? *Journal of Criminal Justice, 34*, 147-152. http://dx.doi.org/10.1016/j.jcrimjus.2006.01.007

Klockars, C. B., Ivković, S. K., & Haberfeld, M. R. (2004). *The contours of police integrity*. Thousand Oaks, CA: Sage.

Lebovich, J., & Rabin, C. (2010, December 17). Officer in fatal shootings arrested after FBI probe. *The Miami Herald*, p. B5.

Lentz, S. A., & Chaires, R. H. (2007). The invention of Peel's principles: A study of policing "textbook" history. *Journal of Criminal Justice, 35*, 65-79. http://dx.doi .org/10.1016/j.jcrimjus.2006.11.016

Lersch, K. M., Bazley, T., & Mieczkowski, T. (2006). Early intervention programs: An effective police accountability tool, or punishment of the productive? *Policing: An International Journal of Police Strategies and Management, 29,* 58-76. http://dx.doi.org/10.1108/13639510610648485

Luo, M., & Dappen, L. (2005). Mixed-methods design for an objective-based evaluation of a magnet school assistance project. *Evaluation and Program Planning, 28,* 109-118. http://dx.doi.org/10.1016/j.evalprogplan.2003.12.001

Macaulay, T. B. M. (1908). *The life and writings of Addison*. New York, NY: Charles Scribner's Sons.

Martinez, P. E. (2008). Projecting prison populations starting with projected admissions. The *Prison Journal, 88,* 493-516. http://dx.doi.org/10.1177/0032885508325395

Mathis, R. L., & Jackson, J. H. (2007). *Human resource management: Essential* perspectives (4th ed.). Mason, OH: Thomson Business and Economics.

McClellan, S. E., & Gustafson, B. G. (2012). Communicating law enforcement professionalization: Social construction of standards. *Policing: An International Journal of Police Strategies and Management, 35*, 104-123. http://dx.doi.org /10.1108/13639511211215478

McDonough, M. E. (2011). The employee wellness plan: A strategy for fighting the "evil from within". *FBI Law Enforcement Bulletin, 80*(12), 1-6.

McElvain, J. P., & Kposowa, A. J. (2004). Police officers characteristics and internal affairs investigations for use of force allegations. *Journal of Criminal Justice, 32*, 265-279. http://dx.doi.org/10.1016/j.jcrimjus.2004.02.006

McMillan, J. H., & Schumacher, S. (2009). *Research in education: Evidence-based inquiry* (6th ed.). Boston, MA: Pearson.

Merriam, S. B. (1998). *Qualitative research and case study applications in education: Revised and expanded from case study research in education.* San Francisco, CA: John Wiley & Sons.

Mertler, C. A., & Charles, C. M. (2010). *Introduction to educational research* (7th ed.). Boston, MA: Pearson.

Miller, J., & Davis, R. C. (2008). Unpacking public attitudes to the police: Contrasting perceptions of misconduct with traditional measures of satisfaction. *International Journal of Police Science and Management, 10*, 9-22. http://dx.doi.org/10.1350/ijps.2008.10.1.9

Miller, S. (2010). Integrity systems and professional reporting in police organizations. *Criminal Justice Ethics, 29*, 241-257. http://dx.doi.org/10.1080/0731129X.2010.524036

Mills, G. E. (2010). *Action research: A guide for the teacher researcher* (4th ed.). Upper Saddle River, NJ: Pearson.

Minor, K. I., Wells, J. B., Angel, E., & Matz, A. K. (2011). Predictors of early job turnover among juvenile correctional facility staff. *Criminal Justice Review, 36,* 58-75. http://dx.doi.org/10.1177/0734016810379253

Nickels, E. L. (2007). A note on the status of discretion in police research. *Journal of Criminal Justice, 35*, 570-578. http://dx.doi.org/10.1016/j.jcrimjus.2007.07.009

Noble, J. J., & Alpert, G. P. (2009). *Managing accountability systems for police conduct: Internal affairs and external oversight*. Long Grove, IL: Waveland Press.

Ono, M., Sachau, D. A., Deal, W. P., Englert, D. R., & Taylor, M. D. (2011). Cognitive ability, emotional intelligence, and the big five personality dimensions as predictors of criminal investigator performance. *Criminal Justice and Behavior, 38*, 471-491. http://dx.doi.org/10.1177/0093854811399406

Orrick, W. D. (2008). *Recruitment, retention, and turnover of police personnel: Reliable, practical, and effective solutions*. Springfield, IL: Charles C. Thomas.

Perrelli, T. (2011, August). *Keynote address.* Presented at the meeting of the Office of Community Oriented Policing Services, Washington, DC.

Pogarsky, G., & Piquero, A. R. (2004). Studying the reach of deterrence: Can deterrence theory help explain police misconduct? *Journal of Criminal Justice, 32*, 371-386. http://dx.doi.org/10.1016/j.jcrimjus.2004.04.007

Police Executive Research Forum. (2010). *Critical issues in policing series*: *Is the economic downturn fundamentally changing how we police?* Washington, DC: Author.

Police Executive Research Forum. (2013). *Civil rights investigations of local police: Lessons learned*. Washington, DC: Author.

Prenzler, T. (2006). Senior police managers' views on integrity testing, and drug and alcohol testing. *Policing: An International Journal of Police Strategies and Management, 29*, 394-407. http://dx.doi.org/10.1108/13639510610684665

Prenzler, T. (2009). *Police corruption: Preventing misconduct and maintaining integrity*. Boca Raton, LA: CRC Press.

Preskill, H. (2008). Evaluation's second act: A spotlight on learning. *American Journal of Evaluation, 29*, 127-138. http://dx.doi.org/10.1177/1098214008316896

Regehr, C., & Bober, T. (2005). *In the line of fire: Trauma in the emergency services*. Oxford, NY: Oxford University Press.

Richards, N. (2010). Police loyalty redux. *Criminal Justice Ethics, 29*, 221-240. http://dx.doi.org/10.1080/0731129X.2010.524033

Ross, D. L., & Bodapati, M. R. (2006). A risk management analysis of the claims, litigation, and losses of Michigan law enforcement agencies: 1985–1999. *Policing: An International Journal of Police Strategies and Management, 29,* 38-57. http:/dx.doi.org/10.1108/13639510610648476

Rossler, M. T., & Terrill, W. (2012). Police responsiveness to service-related requests. *Police Quarterly, 15*, 3-24. http://dx.doi.org/10.1177/1098611111432679

Sagoe, D. (2012). Precincts and prospects in the use of focus groups in social and behavioral science research. *The Qualitative Report, 17*(Article 29), 1-16. Retrieved from http://www.nova.edu/ssss/QR/QR17/sagoe.pdf

Sanders, J. R., & The Joint Committee on Standards for Educational Evaluation. (1994). *The program evaluation standards: How to assess evaluations of educational programs* (2nd ed.). Thousand Oaks, CA: Sage.

Schultz, P. D. (2011, December). *A personnel early warning system for small law enforcement agencies.* Paper presented at the meeting of the International Association of Chiefs of Police, Chicago, IL. Retrieved from http://theiacp.org /Portals/0/pdfs/Implementing_Early_Warning_Handout.pdf

Schwandt, T. A. (2008). Educating for intelligent belief in evaluation. *American Journal of Evaluation, 29*, 139-150. http://dx.doi.org/10.1177/1098214008316889

Schwartz, J. C. (2010). Myths and mechanics of deterrence: The role of lawsuits in law enforcement decisionmaking. *UCLA Law Review, 57,* 1023-1094.

Scriven, M. (1991). *Evaluation thesaurus*. Newbury Park, CA: Sage.

Shane, J. M. (2010). Organizational stressors and police performance. *Journal of Criminal Justice, 38,* 807-818. http://dx.doi.org/10.1016/j.jcrimjus.2010.05.008

Shaw, I., Greene, J. C., & Mark, M. M. (2007). *Handbook of evaluation: Policies, programs and practices*. Thousand Oaks, CA: Sage.

Simmons, K. C. (2008). The politics of policing: Ensuring stakeholder collaboration in the federal reform of local law enforcement agencies. *The Journal of Criminal Law and Criminology, 98,* 489-546. Retrieved from http://scholarlycommons .law.northwestern.edu/jclc/vol98/iss2/

Smith, N. J., & Stoves, D. R. (2012). The DNA of mixed methods: Navigating the proposal and dissertation. In R. Bowden, & E. Ortlieb (Eds.), *Educational research and innovations: 2012 CEDER yearbook* (pp. 177-188). Corpus Christi TX: Texas A&M University Consortium for Educational Development, Evaluation, and Research.

Smith, N. L., Brandon, P. R., Hwalek, M., Kistler, S. J., Labin, S. N., Rugh, J., . . . Yarnall, L. (2011). If not to predict, at least to envision, evaluation's future. *American Journal of Evaluation, 32*, 565-599. http://dx.doi.org/10.1177 /1098214011421412

Spinelli, T. (2010, August). City of North Charleston Police Department: Legitimacy in every action. *Community Policing Dispatch, 3*(8). Retrieved from the Community Oriented Policing Services website: http://cops.usdoj.gov/html/dispatch /August_2010/ncpd.asp

Stinson, P. M., Sr., Liederbach, J., & Freiburger, T. L. (2010). Exit strategy: An exploration of late-stage police crime. *Police Quarterly, 13*, 413-435. http://dx.doi.org/10.1177/1098611110384086

Stewart, D. W., Shamdasani, P. N., & Rook, D. W. (2007). *Focus groups: Theory and practice*. Thousand Oaks, CA: Sage.

Stufflebeam, D. L. (1999, April). *Foundational models for 21st-century program evaluation*. Paper presented at the annual meeting of the American Educational Research Association, Montreal, Quebec, Canada.

Stufflebeam, D. L. (Ed.). (2001). *Evaluation models* (No. 89). San Francisco, CA: Jossey-Bass.

Stufflebeam, D. L. (2003, October). *The CIPP model for evaluation*. Paper presented at the meeting of the Oregon Program Evaluators Network, Portland, OR.

Stufflebeam, D. L. (2007). *CIPP evaluation model checklist* (2nd ed.). Retrieved from the Western Michigan University website: http://www.wmich.edu/evalctr /archive_checklists/cippchecklist_mar07.pdf

Stufflebeam, D. L., & Shinkfield, A. J. (2007). *Evaluation theory, models, and applications*. San Francisco, CA: Jossey-Bass.

SurveyMonkey®. (2014). Create surveys. Get answers. Retrieved from http://www .surveymonkey.com

Temin, P. (2010). *The great recession and the great depression* (NBER Working Paper No. 15645). Cambridge, MA: National Bureau of Economic Research. Retrieved from http://www.nber.org/papers/w15645

Treverton, G. F., Wollman, M., Wilke, E., & Lai, D. (2011). *Moving toward the future of policing*. Santa Monica, CA: Rand.

Tryon, W. W., & Lewis, C. (2009). Evaluating independent proportions for statistical difference, equivalence, indeterminacy, and trivial difference using inferential confidence intervals. *Journal of Educational and Behavioral Statistics, 34,* 171-189. http://dx.doi.org/10.3102/1076998609332753

Tyler, T. R. (2010). Legitimacy and criminal justice: The benefits of self-regulation. *Ohio State Journal of Criminal Law, 7*, 307-359. Retrieved from http://moritzlaw.osu.edu/osjcl/Articles/Volume7_1/Tyler-FinalPDf.pdf

Urban, J. B., & Trochim, W. (2009). The role of evaluation in research–practice integration working toward the "golden spike." *American Journal of Evaluation*, *30*, 538-553. http:dx.doi.org/10.1177/1098214009348327

U.S. Commission on Civil Rights. (2000). *Revisiting who is guarding the guardians? A report on police practices and civil rights in America*. Washington, DC: Author.

U.S. Department of Justice. (2011). *Taking stock: Report from the 2010 roundtable on the state and local law enforcement Police Pattern or Practice Program* (42 USC 14141). Washington, DC: Author.

U.S. Department of Justice, Office of Community Oriented Policing Services. (2009). *Building trust between the police and the citizens they serve: An internal affairs promising practices guide for local law enforcement*. Washington, DC: Author.

U.S. Department of Justice, Office of Public Affairs. (2012a, April). *Five New Orleans police officers sentenced on civil rights and obstruction of justice violations in the Danzinger Bridge shooting case.* Retrieved from http://www.justice.gov/opa/pr/2012/April/12-crt-430.html

U.S. Department of Justice, Office of Public Affairs. (2012b, March). *West Memphis, Ark, police officer convicted of criminal civil rights violation.* Retrieved from http://www.justice.gov/opa/pr/2012/March/12-crt-412.html

U.S. Government Accountability Office. (1997). *Fingerprint-based background checks: Implementation of the National Child Protection Act of 1993*. Washington, DC: Government Printing Office.

U.S. Government Accountability Office. (2012). *Designing evaluations: 2012 revision* (GAO-12-208G). Retrieved from http://www.gao.gov/assets/590/588146.pdf

U.S. Senate. (1994). *Constitution of the United States*. Retrieved from http://www.senate .gov/civics/constitution_item/constitution.htm

Vigneswaran, D. V. (2011). Incident reporting: A technique for studying police corruption. *Policing and Society, 21*, 190-213. http://dx.doi.org/10.1080 /10439463.2010.540658

Walker, S. (2001). *Police accountability: The role of citizen oversight*. Belmont, CA: Wadsworth.

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

Walker, S. (2010). *The police in America: An introduction* (7th ed.). Dubuque, IA: McGraw-Hill.

Walker, S., Alpert, G. P., & Kenney, D. J. (2001). *Early warning systems: Responding to the problem police officer*. Retrieved from the International Association of Chiefs of Police website: http://www.theiacp.org/portals/0/pdfs/early_warning_system .pdf

Walker, S., & Macdonald, M. (2009). An alternative remedy for police misconduct: A model state "pattern or practice" statute. *Civil Rights Law Journal, 19,* 479-552. Retrieved from http://civilrightslawjournal.com/issues/19.479.pdf

Walker, S., Milligan, S. O., & Berke, A. (2005). *Supervision and intervention within early intervention systems: A guide for law enforcement chief executives*. Washington, DC: Police Executive Research Forum.

Walumbwa, F. O., Peterson, S. J., Avolio, B. J., & Hartnell, C. A. (2010). An investigation of the relationships among leader and follower psychological capital, service climate, and job performance. *Personnel Psychology, 63*, 937-963. http://dx.doi.org/10.1111/j.1744-6570.2010.01193.x

Weisburd, D. L., & Neyroud, P. (2011). *Police science: Toward a new paradigm.*. Washington, DC: U.S. Department of Justice. Retrieved from the Harvard Kennedy School, Program in Criminal Justice Policy and Management website: http://www.hks.harvard.edu/var/ezp_site/storage /fckeditor/file/pdfs/centers-programs/programs/criminal-justice/NPIP-Police %20Science-TowardaNewParadigm.pdf

Wilson, J. M., & Heinonen, J. A. (2011). Advancing a police science: Implications from a national survey of police staffing. *Police Quarterly, 14*, 277-297. http://dx.doi.org /10.1177/1098611111414001

Worden, R. E., Kim, M., Harris, C. J., Pratte, M. A., Dorn, S. E., & Hyland, S. S. (2013). Intervention with problem officers: An outcome evaluation of an EIS intervention. *Criminal Justice and Behavior, 40,* 409-437. http://dx.doi.org/10

.1177/0093854812458095

Yarbrough, D. B., Shulha, L. M., Hopson, R. K., & Caruthers, F. A. (2011). *The program evaluation standards: A guide for evaluators and evaluation users* (3rd ed.). Thousand Oaks, CA: Sage.

Appendix A

BlueTeam Software: Illustration Similar to Developer's Dashboard

Appendix A

BlueTeam Software: Illustration Similar to Developer's Dashboard

	Use of Force	Complaint	Vehicle Crash	Pursuit	Overall Threshold
Alpha Squad					
Officer 1	■	■	■	■	■
Officer 2	■	○	○	○	■
Officer 3	■	■	■	■	■
Officer 4	■	■	■	■	■
Bravo Squad					
Officer 5	■	■	■	■	■
Officer 6	■	■	◍	◍	■
Officer 7	■	■	■	■	■
Officer 8	■	■	■	■	■
Charlie Squad					
Officer 9	○	■	◍	■	■
Officer 10	■	■	■	■	■
Officer 11	■	■	■	■	■
Officer 12	■	■	■	■	■

Appendix B

Tracked Categories and Thresholds at the LEA

Appendix B

Tracked Categories and Thresholds at the LEA

Category of tracked data	Threshold: Number. of incidents within 12 months
Overall BlueTeam Program alerts	10
Citizen complaints	4
Agency complaints	4
Employee complaints	4
Inmate complaints	4
Shooting investigation	1
Total internal investigations	4
Use-of-force complaints	4
Domestic violence (personally involved)	1
Chain-of-command referral for conduct	1
Use of force during arrest	6
Taser usage	6
Arrest with charge of resisting without violence	6
Arrest with charge of resisting with violence	6
Failure to appear for court appearance	1
Vehicle pursuit	4
Vehicle accident	4

Appendix C

Focus Group Protocol

Appendix C

Focus Group Protocol

As a member of the administration within the internal affairs component, you were personally involved in the practical and theoretical decision making regarding the implementation and continued operation of the BlueTeam Program (BTP). In light of your comprehensive involvement and knowledge of the BTP, the gathering of your perceptions is of upmost importance to conducting a comprehensive evaluation. During this focus group interview, I will be seeking your perceptions on the implementation and operation of the BTP, including problems, successes, and outcomes, whether intended or not. I will document this session by scripting notes. There will be no information included in my scripting that could disclose your identity. In the interest of the committed confidentiality that I have offered to each of you, please do not discuss any of the content of this meeting once we have concluded the session. I will confirm the accuracy and completeness of my notes by confirming the main ideas of your responses prior to conclusion. In addition, participants will be sent written transcripts of the interview to verify that the information they provided was recorded accurately. This focus group interview, including confirmation of your responses, will take no more than 1 hour to complete.

1. Was the BTP implemented as designed regarding anticipated compatibility and functionality with current agency technology?

2. Was the BTP implemented as designed regarding tracked data categories intended to detect trends of officer misconduct?

3. Was the BTP implemented as designed regarding tracked data categories intended to detect trends of complaints against officers?

4. What were the anticipated and unanticipated events or challenges that took place during the implementation of the BTP?

5. What are your thoughts about the sufficiency of budget for appropriate BTP implementation, particularly in regard to staffing and ancillary resources?

6. What are your thoughts about the sufficiency of workload allocation and time for appropriate BTP implementation?

7. What deficiencies or strengths do you perceive regarding the design of the BTP relative to the sufficiency of training, familiarization with, and introduction of the program?

Appendix D

Survey of Line Supervisors

Appendix D

Survey of Line Supervisors

The purpose of the following survey is to obtain necessary data regarding the evaluation of the BlueTeam Program (BTP) at your agency. In light of your comprehensive involvement and knowledge of the BTP as a line supervisor, the gathering of your perceptions is of upmost importance to conducting a comprehensive evaluation. By administering this survey, I am seeking your perceptions on the implementation and operation of the BTP, including problems, successes, and outcomes, whether intended or not. Your responses and identity are completely confidential, and you may take as much time as needed to complete the survey. Thank you for your time and assistance.

1. What is your understanding of the purpose for the implementation and use of the BTP as originally designed by the LEA administrators?

2. What is your opinion regarding the sufficiency of training and education previously provided by the LEA administrators on the topic of BTP purpose and use for line supervisors?

3. What is your opinion regarding the strengths and weaknesses of the BTP as it was originally designed regarding the implementation and functionality for line supervisors?

4. What is your perception of the BTP regarding the functionality and compatibility with current agency technology such as hardware and software platforms?

5. What is your perception of the BTP functionality regarding the tracked data categories intended to detect trends of officer misconduct?

6. What is your perception of the BTP functionality regarding the tracked data categories intended to detect trends of complaints against officers?

7. What is your perception of the BTP pertaining to ease of use or challenges in overall functionality for line supervisors?

Appendix E

Results

Appendix E

Results

Table E1

Summary of Proportions: 2008 – 2010

	Misconduct complaints		All complaints	
Category	Before BTP[a]	After BTP[b]	Before BTP[c]	After BTP[d]
Complaints	0.130000	0.127000		
Members	0.030000	0.021000	0.229000	0.163000
Workload	0.000077	0.000051	0.000590	0.000401

Note. Total agency members/workload before BTP = 3,176/1,233,873. Total agency members/workload after BTP = 3,846/1,565,282.
[a]Total cases = 95. [b]Total cases = 80. [c]Total cases = 728. . [d]Total cases = 628.

Table E2

Results of Z-Test of Proportions: 2008 – 2010

	Proportion		
Category	Before BTP[a]	After BTP[b]	*z*
Misconduct to total complaints	0.130000	0.127000	0.165
Misconduct to total members	0.030000	0.021000	2.401**
Misconduct to total workload	0.000077	0.000051	2.733**
Total complaints to total members	0.229000	0.163000	3.039**
Total complaints to total workload	0.000590	0.000401	7.135***

Note. N = 918. *z*-test one-tailed, upper.
* $p < 0.05$. ** $p < 0.01$. *** $p < 0.001$.

Table E3:

Descriptive Statistics: 2008 – 2010

Variable	Minimum	Maximum	Mean	*SD*
Workload	92.00	12,864.00	3,716.03	2,374.16
Citizen complaints	00.00	11.00	0.82	1.23
Agency complaints	00.00	3.00	0.10	0.37
Employee complaints	00.00	2.00	0.02	0.17
Inmate complaints	00.00	3.00	0.11	0.37
Use-of-force complaints	00.00	3.00	0.23	0.50
Chain of command	00.00	2.00	0.04	0.23
Response to resistance	00.00	26.00	3.27	4.80
Taser	00.00	9.00	0.66	1.38
Resist arrest without violence	00.00	34.00	4.46	6.48
Resist arrest with violence	00.00	7.00	0.69	1.15
Failure to appear	00.00	2.00	0.12	0.34
Vehicle pursuits	00.00	7.00	0.57	1.01
Vehicle accidents	00.00	4.00	0.90	1.03
Sustained misconduct	00.00	4.00	0.16	0.50
Combined complaints	00.00	12.00	1.27	1.68

$N = 238$

Table E4:

Multiple Regression Model 1: Research Question 6

Independent variable	*B*	*SE*	*b*	*t*
(Constant)	0.04	0.04		0.95
Citizen complaints	0.05	0.03	0.11	1.61
Agency complaints	0.72	0.08	0.54	9.59**
Employee complaints	0.14	0.17	0.05	0.79
Inmate complaints	0.12	0.10	0.09	1.21
Use-of-force complaints	-0.01	0.08	-0.01	-0.15
Chain of command	-0.08	0.13	-0.04	-0.61
Response to resistance	0.01	0.01	0.12	0.01
Taser	-0.06	0.03	-0.17	-2.06*
Resist arrest without violence	-0.01	0.01	-0.08	-0.80
Resist arrest with violence	0.00	0.04	0.00	-0.01
Failure to appear	0.03	0.08	0.02	0.40
Vehicle pursuits	-0.01	0.03	-0.02	-0.33
Vehicle accidents	0.03	0.03	0.06	1.11

$N = 238$. Dependent variable: Sustained misconduct. $R^2 = 0.32$
* $p < 0.05$. ** $p < 0.01$.

Table E5:

Multiple Regression Model 2: Research Question 6

Independent variable	*B*	*SE*	*b*	*t*
(Constant)	0.56	0.14		3.89**
Chain of command	1.01	0.42	0.14	2.43*
Response to resistance	0.06	0.04	0.18	1.61
Taser	-0.10	0.10	-0.08	-1.05
Resist arrest without violence	0.10	0.03	0.39	3.94**
Resist arrest with violence	0.00	0.13	0.00	-0.02
Failure to appear	0.28	0.28	0.06	0.98
Vehicle pursuits	0.07	0.10	0.04	0.67
Vehicle accidents	0.01	0.09	0.01	0.09

$N = 238$. Dependent variable: Total complaints. $R^2 = 0.29$
* $p < 0.05$. ** $p < 0.01$.

Table E6:

Multiple Regression Model 1: Research Question 7

Independent variable	*B*	*SE*	*b*	*t*
(Constant)	0.01	0.06		0.20
Workload	0.00	0.00	0.04	0.68
Citizen complaints	0.05	0.03	0.11	1.60
Agency complaints	0.73	0.08	0.54	9.60**
Employee complaints	0.15	0.17	0.05	0.85
Inmate complaints	0.12	0.10	0.09	1.22
Use-of-force complaints	0.01	0.08	-0.01	-0.13
Chain of command	0.08	0.13	-0.04	-0.63
Response to resistance	0.01	0.01	0.12	1.06
Taser	0.06	0.03	-0.17	-2.06*
Resist arrest without violence	0.01	0.01	-0.09	-0.84
Resist arrest with violence	0.00	0.04	0.00	-0.03
Failure to appear	0.03	0.08	0.02	0.40
Vehicle pursuits	0.01	0.03	-0.02	-0.38
Vehicle accidents	0.03	0.03	0.06	1.10

$N = 238$. Dependent variable: Sustained misconduct. $R^2 = 0.32$
* $p < 0.05$. ** $p < 0.01$.

Table E7:

Multiple Regression Model 2: Research Question 7

Independent variable	*B*	*SE*	*b*	*t*
(Constant)	0.66	0.20		3.27**
Chain of command	1.01	0.42	0.14	2.42*
Response to resistance	0.06	0.04	0.18	1.54
Taser	0.10	0.10	-0.08	-1.04
Resist arrest without violence	0.10	0.03	0.39	3.98**
Resist arrest with violence	0.00	0.13	0.00	-0.01
Failure to appear	0.28	0.28	0.06	0.98
Vehicle pursuits	0.07	0.10	0.04	0.72
Vehicle accidents	0.01	0.09	0.01	0.10
All calls	0.00	0.00	-0.04	-0.72

$N = 238$. Dependent variable: Total complaints. $R^2 = 0.29$
* $p < 0.05$. ** $p < 0.01$.